B160

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/00

A2

(11) International Publication Number: WO 99/58558

(43) International Publication Date: 18 November 1999 (18.11.99)

(21) International Application Number: PCT/US99/10567

(22) International Filing Date: 13 May 1999 (13.05.99)

(30) Priority Data:
60/085,343 13 May 1998 (13.05.98) US
60/098,010 26 August 1998 (26.08.98) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications
US 60/085,343 (CIP)
Filed on 13 May 1998 (13.05.98)
US 60/098,010 (CIP)
Filed on 26 August 1998 (26.08.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). YANG, Junming [CN/US]; 7136 Clarendon Street, San Jose, CA 95129 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: CELL SIGNALING PROTEINS

(57) Abstract

The invention provides human cell signaling proteins (CSIGP) and polynucleotides which identify and encode CSIGP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or prevention disorders associated with expression of CSIGP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

CELL SIGNALING PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of cell signaling proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative and inflammatory disorders.

BACKGROUND OF THE INVENTION

Signal transduction is the process of biochemical events by which cells respond to extracellular signals. Extracellular signals are transduced through a biochemical cascade that begins with the binding of a signal molecule such as a hormone, neurotransmitter, or growth factor, to a cell membrane receptor and ends with the activation of an intracellular target molecule. The process of signal transduction regulates a wide variety of cell functions including cell proliferation, differentiation, and gene transcription.

Signal transduction is the general process by which cells respond to extracellular signals (hormones, neurotransmitters, growth and differentiation factors, etc.) through a cascade of biochemical reactions that begins with the binding of the signaling molecule to a cell membrane receptor and ends with the activation of an intracellular target molecule. Intermediate steps in this process involve the activation of various cytoplasmic proteins by phosphorylation via protein kinases and the eventual translocation of some of these activated proteins to the cell nucleus where the transcription of specific genes is triggered. Thus, the signal transduction process regulates all types of cell functions including cell proliferation, differentiation, and gene transcription.

Protein kinases play a key role in the signal transduction process by phosphorylating and activating various proteins involved in signaling pathways. The high energy phosphate which drives this activation is generally transferred from adenosine triphosphate molecules (ATP) to a particular protein by protein kinases and removed from that protein by protein phosphatases. Phosphorylation occurs in response to extracellular signals, cell cycle checkpoints, and environmental or nutritional stresses. Protein kinases are roughly divided into two groups; those that phosphorylate tyrosine residues (protein tyrosine kinases, PTK) and those that phosphorylate serine or threonine residues (serine/threonine kinases, STK). A few protein kinases have dual specificity for serine/threonine and tyrosine residues. Almost all kinases contain a similar 250-300 amino acid catalytic domain containing specific residues and sequence motifs characteristic of the kinase family. (Hardie, G. and Hanks, S. (1995) The Protein Kinase Facts Books, Vol I:7-20

Academic Press, San Diego, CA.)

STKs include the second messenger dependent protein kinases such as the cyclic-AMP dependent protein kinases (PKA), which are involved in mediating hormone-induced cellular responses; calcium-calmodulin (CaM) dependent protein kinases, which are involved in regulation of smooth muscle contraction, glycogen breakdown, and neurotransmission; and the mitogen-activated protein kinases (MAP) which mediate signal transduction from the cell surface to the nucleus via phosphorylation cascades. Altered PKA expression is implicated in a variety of disorders and diseases including cancer, thyroid disorders, diabetes, atherosclerosis, and cardiovascular disease. (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York, NY, pp. 416-431, 1887.)

PTKs are divided into transmembrane, receptor PTKs and nontransmembrane, non-receptor PTKs. Transmembrane protein-tyrosine kinases are receptors for most growth factors which include epidermal GF, platelet-derived GF, fibroblast GF, hepatocyte GF, insulin and insulin-like GFs, nerve GF, vascular endothelial GF, and macrophage colony stimulating factor. Non-receptor PTKs lack transmembrane regions and, instead, form complexes with the intracellular regions of cell surface receptors. Receptors that function through non-receptor PTKs include those for cytokines, hormones (growth hormone and prolactin) and antigen-specific receptors on T and B lymphocytes.

Many of these PTKs were first identified as the products of mutant oncogenes in cancer cells where their activation was no longer subject to normal cellular controls. In fact, about one third of the known oncogenes encode PTKs, and it is well known that cellular transformation (oncogenesis) is often accompanied by increased tyrosine phosphorylation activity. (Charbonneau H and Tonks NK (1992) Annu Rev Cell Biol 8:463-493.)

Protein phosphatases regulate the effects of protein kinases by removing phosphate groups from molecules previously activated by kinases. The two principle categories of protein phosphatases are the protein phosphatases (PPs) and the protein tyrosine phosphatases (PTPs). PPs dephosphorylate phosphoserine/threonine residues and are important regulators of many cAMP-mediated hormone responses in cells. (Cohen, P. (1989) Annu. Rev. Biochem. 58:453-508.) PTPs reverse the effects of protein tyrosine kinases and play a significant role in cell cycle and cell signaling processes. (Charbonneau and Tonks, supra.) In the process of cell division, for example, a specific PTP (M-phase inducer phosphatase) plays a key role in the induction of mitosis by dephosphorylating and activating a specific PTK (CDC2) leading to cell division. (Sadu, K.. et al. (1990) Proc. Natl. Acad. Sci. 87:5139-5143.)

Guanine nucleotide binding proteins (GTP-binding proteins) are critical mediators of the signal transduction pathway. Extracellular ligands such as hormones, growth factors,

neuromodulators, or other signaling molecules bind to transmembrane receptors, and the signal is propagated to effector molecules by intracellular signal transducing proteins. Many of these signal transduction proteins are GTP-binding proteins which regulate intracellular signaling pathways. GTP-binding proteins participate in a wide range of other regulatory functions including metabolism, growth, differentiation, cytoskeletal organization, and intracellular vesicle transport and secretion. Exchange of bound GDP for GTP followed by hydrolysis of GTP to GDP provides the energy that enables GTP-binding proteins to alter their conformation and interact with other cellular components. Two structurally distinct classes of GTP-binding proteins are recognized: heterotrimeric GTP-binding proteins, consisting of three different subunits, and monomeric, low molecular weight (LMW), GTP-binding proteins consisting of a single polypeptide chain.

G protein coupled receptors (GPCR) are a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines, mediators of inflammation, peptide hormones, and sensory signal mediators. A GPCR becomes activated when the receptor binds to its extracellular ligand. The beta subunit of the GPCR, which consists of an amino-terminal helical segment followed by seven WD, or β transducin repeats, transduces signals across the plasma membrane. Conformational changes in the GPCR, resulting from the ligand-receptor interaction, promote the binding of GTP to the GPCR intracellular domains. GTP binding to the GPCR leads to the interaction of the GPCR alpha subunit with adenylate cyclase or other second messenger molecule generators. This interaction regulates the activity of second messenger molecules such as cAMP, cGMP, or eicosinoids which, in turn, regulate phosphorylation and activation of other intracellular proteins. The GPCR changes conformation upon hydrolysis of the bound GTP by GTPases, dissociates from the second messenger molecule generator, and returns to its initial pre-ligand binding conformation.

G beta proteins, also known as β transducins, contain seven tandem repeats of the WD-repeat sequence motif, a motif found in many proteins with regulatory functions. WD-repeat proteins contain from four to eight copies of a loosely conserved repeat of approximately 40 amino acids which participates in protein-protein interactions. Mutations and variant expression of β transducin proteins are linked with various disorders. Mutations in LIS1, a subunit of the human platelet activating factor acetylhydrolase, cause Miller-Dieker lissencephaly. RACK1 binds activated protein kinase C, and RbAp48 binds retinoblastoma protein. CstF is required for polyadenylation of mammalian pre-mRNA in vitro and associates with subunits of cleavage-stimulating factor. CD4, an integral membrane glycoprotein which functions as an HIV co-receptor for infection of human host cells is degraded by HIV-encoded Vpu in the endoplasmic reticulum. WD repeats of human beta TrCP molecule mediate the formation of the CD4- Vpu, inducing CD4 proteolysis (Neer, E.J. et al. (1994) Nature 371:297-300 and Margottin, F. et al.

(1998) Mol. Cell. 1:565-574).

Irregularities in the GPCR signaling cascade may result in abnormal activation of leukocytes and lymphocytes, leading to the tissue damage and destruction seen in many inflammatory and autoimmune diseases such as rheumatoid arthritis, biliary cirrhosis, hemolytic anemia, lupus erythematosus, and thyroiditis. Abnormal cell proliferation, including cyclic AMP stimulation of brain, thyroid, adrenal, and gonadal tissue proliferation is regulated by G proteins. Mutations in G_α subunits have been found in growth-hormone-secreting pituitary somatotroph tumors, hyperfunctioning thyroid adenomas, and ovarian and adrenal neoplasms (Meij, J.T.A. (1996) Mol. Cell. Biochem. 157:31-38; Aussel, C. et al. (1988) J. Immunol. 140:215-220).

LMW GTP-binding proteins regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. They consist of single polypeptides which, like the alpha subunit of the heterotrimeric GTP-binding proteins, are able to bind to and hydrolyze GTP, thus cycling between an inactive and an active state. LMW GTP-binding proteins respond to extracellular signals from receptors and activating proteins by transducing mitogenic signals involved in various cell functions. The binding and hydrolysis of GTP regulates the response of LMW GTP-binding proteins and acts as an energy source during this process (Bokoch, G. M. and Der, C. J. (1993) FASEB J. 7:750-759).

At least sixty members of the LMW GTP-binding protein superfamily have been identified and are currently grouped into the four subfamilies of ras, rho, arf, sar1, ran, and rab. Activated ras genes were initially found in human cancers and subsequent studies confirmed that ras function is critical in determining whether cells continue to grow or become differentiated. Other members of the LMW G-protein superfamily have roles in signal transduction that vary with the function of the activated genes and the locations of the GTP-binding proteins that initiate the activity. Rho GTP-binding proteins control signal transduction pathways that link growth factor receptors to actin polymerization, which is necessary for normal cellular growth and division. The rab, arf, and sar1 families of proteins control the translocation of vesicles to and from membranes for protein localization, protein processing, and secretion. Ran GTP-binding proteins are located in the nucleus of cells and have a key role in nuclear protein import, the control of DNA synthesis, and cell-cycle progression (Hall, A. (1990) Science 249:635-640; Barbacid, M. (1987) Ann. Rev Biochem. 56:779-827; and Sasaki, T. and Takai, Y. (1998) Biochem. Biophys. Res. Commun. 245:641-645).

LMW GTP-binding proteins are GTPases which cycle between a GTP-bound active form and a GDP-bound inactive form. This cycle is regulated by proteins that affect GDP dissociation, GTP association, or the rate of GTP hydrolysis. Proteins affecting GDP association are

represented by guanine nucleotide dissociation inhibitors and guanine nucleotide exchange factors (GEP). The best characterized is the mammalian homologue of the Drosophila Son-of-Sevenless protein. Proteins affecting GTP hydrolysis are exemplified by GTPase-activating proteins (GAP). Both GEP and GAP activity may be controlled in response to extracellular stimuli and modulated by accessory proteins such as RalBP1 and POB1. The GDP-bound form is converted to the GTP-bound form through a GDP/GTP exchange reaction facilitated by guanine nucleotide-releasing factors. The GTP-bound form is converted to the GDP-bound form by intrinsic GTPase activity, and the conversion is accelerated by GAP (Ikeda, M. et al. (1998) J. Biol. Chem. 273:814-821;Quilliam, L. A. (1995) Bioessays 17:395-404.). Mutant Ras-family proteins, which bind but can not hydrolyze GTP, are permanently activated, and cause cell proliferation or cancer, as do GEP that activate LMW GTP-binding proteins (Drivas, G. T. et al. (1990) Mol. Cell. Biol. 10:1793-1798; and Whitehead, I. P. et al. (1998) Mol Cell Biol. 18:4689-4697.)

The discovery of new cell signaling proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative and inflammatory disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, cell signaling proteins, referred to collectively as "CSIGP" and individually as CSIGP-1 through CSIGP-13 . In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-13, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino

acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-13, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of CSIGP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of CSIGP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO), clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding CSIGP.

Table 2 shows features of each polypeptide sequence including potential motifs, homologous sequences, and methods and algorithms used for identification of CSIGP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis, diseases, disorders or conditions associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which Incyte cDNA clones encoding CSIGP were isolated.

Table 5 shows the programs, their descriptions, references, and threshold parameters used to analyze CSIGP.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and

methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"CSIGP" refers to the amino acid sequences of substantially purified CSIGP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to CSIGP, increases or prolongs the duration of the effect of CSIGP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of CSIGP.

An "allelic variant" is an alternative form of the gene encoding CSIGP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding CSIGP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as CSIGP or a polypeptide with at least one functional characteristic of CSIGP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding CSIGP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding CSIGP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent CSIGP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of CSIGP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and

phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of CSIGP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of CSIGP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to CSIGP, decreases the amount or the duration of the effect of the biological or immunological activity of CSIGP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of CSIGP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind CSIGP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form

duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic CSIGP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding CSIGP or fragments of CSIGP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence"refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding CSIGP, by northern analysis is indicative of the presence of nucleic acids encoding CSIGP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding CSIGP.

A "deletion"refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI). The MEGALIGN program can create alignments between two or more sequences according to different methods, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid

sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of CSIGP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of CSIGP.

The phrases "nucleic acid" or "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may

represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding CSIGP, or fragments thereof, or CSIGP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other

conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of CSIGP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to CSIGP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or

lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human cell signaling proteins (CSIGP), the polynucleotides encoding CSIGP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative and inflammatory disorders.

Table 1 lists the Incyte Clones used to derive full length nucleotide sequences encoding CSIGP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NO) of the amino acid and nucleic acid sequences, respectively. Column 3 shows the Clone ID of the Incyte Clone in which nucleic acids encoding each CSIGP were first identified, and column 4, the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones, their corresponding cDNA libraries, and shotgun sequences useful as fragments in hybridization technologies, and which are part of the consensus nucleotide sequence of each CSIGP.

The columns of Table 2 show various properties of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3, potential phosphorylation sites; column 4, potential glycosylation sites; column 5, the amino acid residues comprising signature sequences and motifs; column 6, homologous sequences; and column 7, analytical methods used to identify each protein through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and disease-association of nucleotide sequences encoding CSIGP. The first column of Table 3 lists the polynuclceotide sequence identifiers. The second column lists tissue categories which express CSIGP as a fraction of total tissue categories expressing CSIGP. The third column lists diseases, disorders, and conditiond associated with those tissues expressing CSIGP. The fourth column lists the vectors used to subclone the cDNA library.

The following fragments of the nucleotide sequences encoding CSIGP are useful in hybridization or amplification technologies to identify SEQ ID NO:14-26 and to distinguish between SEQ ID NO:14-26 and similar polynucleotide sequences. The useful fragments are the

fragment of SEQ ID NO:14 from about nucleotide 135 to about nucleotide 189, the fragment of SEQ ID NO:15 from about nucleotide 493 to about nucleotide 558, the fragment of SEQ ID NO:16 from about nucleotide 1170 to about nucleotide 1233, the fragment of SEQ ID NO:17 from about nucleotide 939 to about nucleotide 996, the fragment of SEQ ID NO:18 from about nucleotide 424 to about nucleotide 486, the fragment of SEQ ID NO:19 from about nucleotide 274 to about nucleotide 333, and the fragment of SEQ ID NO:20 from about nucleotide 1013 to about nucleotide 1070, the fragment of SEQ ID NO:21 from about nucleotide 284 to about nucleotide 325, the fragment of SEQ ID NO:22 from about nucleotide 642 to about nucleotide 674, the fragment of SEQ ID NO:23 from about nucleotide 742 to about nucleotide 769, the fragment of SEQ ID NO:24 from about nucleotide 457 to about nucleotide 486, the fragment of SEQ ID NO:25 from about nucleotide 205 to about nucleotide 246, and the fragment of SEQ ID NO:26 from about nucleotide 319 to about nucleotide 342.

The invention also encompasses CSIGP variants. A preferred CSIGP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the CSIGP amino acid sequence, and which contains at least one functional or structural characteristic of CSIGP.

The invention also encompasses polynucleotides which encode CSIGP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:14-26 which encodes CSIGP.

The invention also encompasses a variant of a polynucleotide sequence encoding CSIGP. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding CSIGP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:14-26 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:14-26. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of CSIGP

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding CSIGP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide

sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring CSIGP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode CSIGP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring CSIGP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding CSIGP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding CSIGP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode CSIGP and CSIGP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding CSIGP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:14-26 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment,

hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the HYDRA microdispenser (Robbins Scientific, Sunnyvale CA), MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA Sequencing Systems (Perkin-Elmer) or the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA). The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding CSIGP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect

upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode CSIGP may be cloned in recombinant DNA molecules that direct expression of

CSIGP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express CSIGP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter CSIGP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding CSIGP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.) Alternatively, CSIGP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of CSIGP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active CSIGP, the nucleotide sequences encoding CSIGP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding CSIGP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding CSIGP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding CSIGP and its initiation

codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding CSIGP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding CSIGP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding CSIGP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding CSIGP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding CSIGP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of CSIGP are needed, e.g. for the production of antibodies, vectors which direct high level expression of CSIGP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage

promoter may be used.

Yeast expression systems may be used for production of CSIGP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Grant et al. (1987) Methods Enzymol. 153:516-54; and Scorer, C. A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of CSIGP. Transcription of sequences encoding CSIGP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding CSIGP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses CSIGP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of CSIGP in cell lines is preferred. For example, sequences encoding CSIGP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2

days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ or *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding CSIGP is inserted within a marker gene sequence, transformed cells containing sequences encoding CSIGP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding CSIGP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding CSIGP and that express CSIGP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of CSIGP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques

include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on CSIGP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols.Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding CSIGP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding CSIGP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding CSIGP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode CSIGP may be designed to contain signal sequences which direct secretion of CSIGP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK. HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Bethesda MD) and may be chosen to ensure the

correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding CSIGP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric CSIGP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of CSIGP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the CSIGP encoding sequence and the heterologous protein sequence, so that CSIGP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled CSIGP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of CSIGP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of CSIGP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between CSIGP and cell signaling proteins. In addition, the expression of CSIGP is closely associated with cell proliferation and inflammatory disorders. Therefore, in cell proliferative and

inflammatory disorders where CSIGP is an inhibitor or suppressor of cell proliferation, it is desirable to increase the expression of CSIGP. In cell proliferative and inflammatory disorders where CSIGP is an activator or enhancer and is promoting cell proliferation, it is desirable to decrease the expression of CSIGP.

Therefore, in one embodiment, CSIGP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CSIGP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma.

In another embodiment, a vector capable of expressing CSIGP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CSIGP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified CSIGP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CSIGP including, but not limited to, those provided above.

In still another embodiment. an agonist which modulates the activity of CSIGP may be

administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CSIGP including, but not limited to, those listed above.

In a further embodiment, an antagonist of CSIGP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of CSIGP. Examples of such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds CSIGP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express CSIGP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding CSIGP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of CSIGP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of CSIGP may be produced using methods which are generally known in the art. In particular, purified CSIGP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind CSIGP. Antibodies to CSIGP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with CSIGP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to

CSIGP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of CSIGP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to CSIGP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce CSIGP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for CSIGP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between

CSIGP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering CSIGP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for ABBR. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of ABBR-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple ABBR epitopes, represents the average affinity, or avidity, of the antibodies for ABBR. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular ABBR epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the ABBR-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of ABBR, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of ABBR-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding CSIGP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding CSIGP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding CSIGP. Thus, complementary molecules or fragments may be used to modulate CSIGP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding CSIGP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses,

or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding CSIGP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding CSIGP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding CSIGP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding CSIGP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding CSIGP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be

prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding CSIGP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of CSIGP, antibodies to CSIGP, and mimetics, agonists, antagonists, or inhibitors of CSIGP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal,

enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl

cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of CSIGP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example CSIGP or fragments thereof, antibodies of CSIGP, and agonists, antagonists or inhibitors of CSIGP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or

LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind CSIGP may be used for the diagnosis of cell proliferative and inflammatory disorders characterized by expression of CSIGP, or in assays to monitor patients being treated with CSIGP or agonists, antagonists, or inhibitors of CSIGP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for CSIGP include methods which utilize the antibody and a label to detect CSIGP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring CSIGP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of CSIGP expression. Normal or standard values for CSIGP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to

CSIGP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of CSIGP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding CSIGP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of CSIGP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of CSIGP, and to monitor regulation of CSIGP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding CSIGP or closely related molecules may be used to identify nucleic acid sequences which encode CSIGP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding CSIGP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the CSIGP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:14-26 or from genomic sequences including promoters, enhancers, and introns of the CSIGP gene.

Means for producing specific hybridization probes for DNAs encoding CSIGP include the cloning of polynucleotide sequences encoding CSIGP or CSIGP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding CSIGP may be used for the diagnosis of cell proliferative and inflammatory disorders associated with expression of CSIGP. Examples of such disorders include, but are not limited to, a disorder of cell proliferation such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease

(MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The polynucleotide sequences encoding CSIGP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and ELISA assays; and in microarrays utilizing fluids or tissues from patients to detect altered CSIGP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding CSIGP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding CSIGP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding CSIGP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of CSIGP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a

sequence, or a fragment thereof, encoding CSIGP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding CSIGP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding CSIGP, or a fragment of a polynucleotide complementary to the polynucleotide encoding CSIGP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of CSIGP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to

determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding CSIGP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding CSIGP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, CSIGP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between CSIGP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with CSIGP, or fragments thereof, and washed. Bound CSIGP is then detected by methods well known in the art. Purified CSIGP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding CSIGP specifically compete with a test compound for binding CSIGP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with CSIGP.

In additional embodiments, the nucleotide sequences which encode CSIGP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any was whatsoever.

The entire disclosure of all applications, patents, and publications, cited above and below, and of US provisional applications 60/085,343 (filed May 13, 1998), and 60/098,010 (filed August 26, 1998) are hereby incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated

from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Valencia CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the REAL Prep 96 plasmid kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II

fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using either an ABI CATALYST 800 (Perkin-Elmer) or a HYDRA microdispenser (Robbins) or MICROLAB 2200 (Hamilton) sequencing preparation system in combination with PTC-200 thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems of the MEGABACE 1000 DNA sequencing system (Molecular Dynamics) and ABI protocols, base calling software, and kits (Perkin-Elmer). Alternatively, solutions and dyes from Amersham Pharmacia Biotech were used. Reading frames were determined using standard methods (Ausubel, 1997, supra). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the software programs, descriptions, references, and threshold parameters used. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides a brief description thereof, the third column presents the references which are incorporated by reference herein, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the probability the greater the homology). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, S. San Francisco CA) and LASERGENE software (DNASTAR).

cDNAs were also compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 sequence analysis system. In this algorithm, Pattern Specification Language (TRW Inc, Los Angeles, CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT- 670 sequence analysis system using the methods similar to those used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance

matches.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:14-26. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported a percentage distribution of libraries in which

the transcript encoding CSIGP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease or condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease expression are reported in Table 3.

V. Extension of CSIGP Encoding Polynucleotides

The full length nucleic acid sequence of SEQ ID NO:14-26 was produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture

was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequence of SEQ ID NO:14-26 is used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Choice, Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:14-26 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following

endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba1, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the blots to film for several hours, hybridization patterns are compared visually.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the CSIGP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring CSIGP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of CSIGP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit

translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the CSIGP-encoding transcript.

IX. Expression of CSIGP

Expression and purification of CSIGP is achieved using bacterial or virus-based expression systems. For expression of CSIGP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express CSIGP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of CSIGP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding CSIGP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, CSIGP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from CSIGP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified CSIGP obtained by these methods can be used directly in the following activity assay.

X. Dem nstration f CSIGP Activity

CSIGP activity can be assayed in vitro by monitoring the mobilization of Ca^{++} as part of the signal transduction pathway. (See, e.g., Grynkievwicz, G. et al. (1985) J. Biol. Chem.

260:3440; McColl, S. et al. (1993) J. Immunol. 150:4550-4555; and Aussel, C. et al. (1988) supra) The assay requires preloading neutrophils or T cells with a fluorescent dye such as FURA-2 or BCECF (Universal Imaging Corp, Westchester PA) whose emission characteristics have been altered by Ca^{++} binding. When the cells are exposed to one or more activating stimuli artificially (ie, anti-CD3 antibody ligation of the T cell receptor) or physiologically (ie, by allogeneic stimulation), Ca^{++} flux takes place. This flux can be observed and quantified by assaying the cells in a fluorometer or fluorescent activated cell sorter. Measurements of Ca^{++} flux are compared between cells in their normal state and those preloaded with CSIGP.

Protein kinase activity in CSIGP is determined by measuring the phosphorylation of a protein substrate using gamma-labeled ^{32}P-ATP and quantitation of the incorporated radioactivity using a radioisotope counter. CSIGP is incubated with the protein substrate, ^{32}P-ATP, and an appropriate kinase buffer. The ^{32}P incorporated into the product is separated from free ^{32}P-ATP by electrophoresis and the incorporated ^{32}P is counted. The amount of ^{32}P recovered is proportional to the activity of CSIGP in the assay. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

Protein phosphatase (PP) activity in CSIGP is determined by measuring the hydrolysis of P-nitrophenyl phosphate (PNPP). CSIGP is incubated together with PNPP in HEPES buffer pH 7.5, in the presence of 0.1% b-mercaptoethanol at 37°C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH and the increase in light absorbance at 410 nm resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the activity of CSIGP in the assay.

XI. Production of CSIGP Specific Antibodies

CSIGP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the CSIGP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for

antipeptide activity by, for example, binding the peptide to plastic. blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XII. Purification of Naturally Occurring CSIGP Using Specific Antibodies

Naturally occurring or recombinant CSIGP is substantially purified by immunoaffinity chromatography using antibodies specific for CSIGP. An immunoaffinity column is constructed by covalently coupling anti-CSIGP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing CSIGP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of CSIGP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/CSIGP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and CSIGP is collected.

XIII. Identification of Molecules Which Interact with CSIGP

CSIGP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled CSIGP, washed, and any wells with labeled CSIGP complex are assayed. Data obtained using different concentrations of CSIGP are used to calculate values for the number, affinity, and association of CSIGP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	14	016108	HUVELPB01	016108, 016624, (HUVELPB01), 970134 (MUSCNOT02), 1605858 (LUNGNOT15), 1419046 (KIDNNOT09)
2	15	640521	BRSTNOT03	640521 (BRSTNOT03)
3	16	1250171	LUNGFET03	1250171 (LUNGFET03), 260744 (HNT2RAT01), 077085 (SYNORAB01), 2790184 (COLNTUT16), SAEB01398, SAEB00499, SAEB02190, SAEB00648, SAEB00948
4	17	1911587	CONNTUT01	1911587 (CONNTUT01), 1989659 (CORPNOT02)
5	18	2079081	ISLTNOT01	2079081 (ISLTNOT01), 2631449 (COLNTUT15), 2350624 (COLSUCT01), 2568459 (HIPOAZT01), 2132860 (OVARNOT03)
6	19	2472655	THP1NOT03	2472655 (THP1NOT03), 1325950 (LPARNOT02), SAEA01014, SAEA01114, SAEA03382
7	20	2948818	KIDNFET01	2948818 (KIDNFET01), 1543592 (PROSTUT04), SAAE00176

Table 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
8	21	054191	FIBRNOT01	054191H1 and 054191R6 (FIBRNOT01), 483547H1, 483547R6, and 483547T6 (HNT2RAT01), 1537974R6 (SINTTUT01), 1633493H1 (COLNNOT19)
9	22	1403604	LATRTUT02	491348H1 (HNT2AGT01), 1403604H1 (LATRTUT02), 3331135T6.com (BRAIFET01), SBAA02561F1.comp, SBAA03200F1, SBAA01960F1.comp, SBAA01439F1, SBAA01304F1
10	23	1652936	PROSTUT08	467767R6 (LATRNOT01), 1551938R6 (PROSNOT06), 1652936F6 and 1652936H1 (PROSTUT08), 1817388F6 and 1817388H1 (PROSNOT20), 2822521H1 (ADRETUT06)
11	24	1710702	PROSNOT16	1474380T1 (LUNGTUT03), 1710702H1 (PROSNOT16), 2189187H1 (PROSNOT26), 1526267F1 (UCMCL5T01), 1467104F1 (PANCTUT02)
12	25	3239149	COLAUCT01	482693H1 (HNT2RAT01), 2287788R6 (BRAINON01), 2570350T6 (HIPOAZT01), 3239149F6 and 3239149H1 (COLAUCT01), 3837574F6 (DENDTNT01), 4993747H1 (LIVRTUT11)
13	26	3315936	PROSBPT03	2501356T6 (ADRETUT05), 3315936H1 (PROSBPT03)

Table 2

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Homologous Sequence	Analytical Methods
1	418	S359 S2 T12 S56 T91 T257 S287 S306 T402 S414 T9 S16 S43 T87 S184 S327 S334	N54 N70 N118	Y58-I293	Serine /threonine protein kinase	BLOCKS PRINTS PFAM
2	540	S100 T145 S26 T56 S100 T166 S358 S456 T462 T467 S503 S11 S30 S95 S137 S197 T280 T362 S367 S474 Y234 Y305	N460	Y165-V446	Ca2 +/calmodulin-dependent protein kinase kinase	BLOCKS PRINTS MOTIFS BLAST PFAM
3	729	T96 S348 T373 S518 S531 T682 T78 T239 T478 Y235	N42 N455 N614	W9-I238	Serine/ threonine protein kinase	BLOCKS PFAM PRINTS MOTIFS BLAST
4	313	S38 S82 S95 S97 T143 Y30	N79 N80 N172 N192	R114-S135	Protein tyrosine phosphatase	PRINTS BLAST

Table 2 cont.

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Homologous Sequence	Analytical Methods
5	506	S114 S300 S81 S160 T162 S211 S253 S291 S335 S341 T63 S143 T144 S156 T177 S196 S363 S439 Y45 Y187	N275	SH3 domains: R441-L495	PEST phosphatase interacting protein	BLOCKS PRINTS PFAM BLAST
6	341	S39 S118 T125 S180 S110 S170 S173 S195 T299	N37 N178 N229 N263		Prolactin receptor associated protein (PRAP)	BLAST
7	898	S56 T640 S15 S107 T210 T267 S324 S366 S374 S504 T547 T592 T640 S655 T681 T756 S775 S58 S249 T437 S551 T573 S655 T726 T745 T762 S836 S858 S879	N322 N347 N389 N502 N503	F24-V277	Serine/ threonine protein kinase	BLOCKS PRINTS PFAM MOTIFS BLAST

Table 2 cont.

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Homologous Sequence	Analytical Methods
8	336	S34 T110 S148 S311	N137 N144 N169	T175-I195 V236-T254	putative G-protein-coupled receptor	PRINTS, BLAST HMM, Motifs
9	686	T192 S312 S483 S502 S23 T584	N17 N457 N618 N642	G544-N560	GDP-GTP exchange protein	PRINTS, BLAST Motifs
10	519	S3 S77 S130 S176 S187 T196 S245 S265 T280 T290 T305 T324 S325 S351 S384 S390 T29 S33 S265 T305 S311 T453 S464 Y131 Y145	N128		GTPase-interacting protein	BLAST Motifs
11	334	S332 T186 S198 S269 T321 S90 S139 Y289	N20 N30	L267-L281	G-protein beta WD-40 repeat containing protein	PRINTS, BLAST Motifs
12	569	S91 S19 S109 S162 S376 S418 T514 S535 S536 S19 S39 T266 T288 T328 T381 T411 T451 S519	N17 N77 N416	L320-V334 M360-M374 I403-T417 V443-I457 I483-L497 I532-F546	beta-transducin repeats containing protein	PRINTS, BLAST PFAM, Motifs
13	123	S14 T107 Y44 Y70	N100	M1-N52	SAR1 family GTP-binding protein	PRINTS, BLOCKS BLAST, Motifs

Table 3

Polynuleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
14	Cardiovascular (0.194) Hematopoietic/Immune (0.194) Developmental (0.139)	Cancer (0.389) Inflammation (0.333) Cell proliferative (0.306)	pBLUESCRIPT
15	Reproductive (0.282) Nervous (0.179) Developmental (0.128)	Cancer (0.410) Cell proliferative (0.205) Inflammation (0.154)	pSPORT1
16	Reproductive (0.286) Hematopoietic/Immune (0.167) Nervous (0.119)	Cancer (0.429) Inflammation (0.310) Cell proliferative (0.214)	pINCY
17	Nervous (0.235) Reproductive (0.147) Gastrointestinal (0.118)	Cancer (0.471) Cell proliferative (0.176) Trauma (0.176)	pINCY
18	Reproductive (0.400) Gastrointestinal (0.267) Cardiovascular (0.133)	Cancer (0.533) Inflammation (0.333) Cell proliferative (0.067)	pINCY
19	Nervous (0.273) Hematopoietic/Immune (0.227) Reproductive (0.227)	Cancer (0.364) Inflammation (0.364) Cell proliferative (0.318)	pINCY
20	Hematopoietic/Immune (0.216) Reproductive (0.216) ervous (0.157)	Cancer (0.412) Inflammation (0.294) Cell proliferative (0.216)	pINCY

Table 3 cont.

Polynucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
21	Cardiovascular (0.217) Gastrointestinal (0.174) Nervous (0.174)	Cell proliferative (0.652) Inflammation (0.304)	pBlUESCRIPT
22	Reproductive (0.370) Nervous (0.222) Hematopoietic/Immune (0.148)	Cell proliferative (0.778) Trauma (0.148)	pINCY
23	Reproductive (0.400) Cardiovascular (0.200) Hematopoietic/Immune (0.133)	Cancer (0.533) Inflammation (0.200)	pINCY
24	Reproductive (0.241) Nervous (0.190) Cardiovascular (0.138)	Cell proliferative (0.724) Inflammation (0.138)	pINCY
25	Musculoskeletal (0.222) Nervous (0.222) Gastrointestinal (0.167)	Cell proliferative (0.555) Inflammation (0.222)	pINCY
26	Reproductive (0.750) Cardiovascular (0.250)	Cancer (0.500) Inflammation (0.500)	pINCY

0

Table 4

Polynucleotide SEQ ID NO:	Library	Library Description
14	HUVELPB01	The library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells that were stimulated with cytokine/LPS. HUV-EC-C is an endothelial cell line derived from the vein of a normal human umbilical cord. RNA was isolated from two pools of HUV-EC-C cells that had been treated with either gamma IFN and TNF-alpha or IL-1 beta and LPS.
15	BRSTNOT03	The library was constructed using RNA isolated from nontumorous breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma. Family history included benign hypertension, hyperlipidemia, and a malignant neoplasm of the colon.
16	LUNGFET03	The library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation from fetal demise. Family history included bronchitis.
17	CONNTUT01	The library was constructed using RNA isolated from a soft tissue tumor removed from the clival area of the skull of a 30-year-old Caucasian female. Pathology indicated chondroid chordoma with neoplastic cells reactive for keratin. Patient history included deficiency anemia.
18	ISLTNOT01	The library was constructed using RNA isolated from pancreatic islet cells. Starting RNA was made from a pooled collection of islet cells.
19	THP1NOT03	The library was constructed using RNA isolated from untreated THP-1 cells. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia.
20	KIDNFET01	The library was constructed using RNA isolated from kidney tissue removed from a Caucasian female fetus, who died at 17 weeks' gestation from ancephalus. Family history included gastritis.

Table 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Description
21	FIBRNOT01	The library was constructed at Stratagene (STR937212), using RNA isolated from the WI38 lung fibroblast cell line, which was derived from a 3-month-old Caucasian female fetus. 2x10e6 primary clones were amplified to stabilize the library for long-term storage.
22	LATRTUT02	The library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease and hyperlipidemia. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
23	PROSTUT08	The library was constructed using RNA isolated from prostate tumor tissue removed from a 60-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included tuberculosis, cerebrovascular disease, and arteriosclerotic coronary artery disease.
24	PROSNOT16	The library was constructed using RNA isolated from diseased prostate tissue removed from a 68-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+4). The patient presented with elevated prostate specific antigen (PSA). During this hospitalization, the patient was diagnosed with myasthenia gravis. Patient history included osteoarthritis, and type II diabetes. Family history included benign hypertension, acute myocardial infarction, hyperlipidemia, and arteriosclerotic coronary artery disease.

Table 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Description
25	COLAUCT01	The library was constructed using RNA isolated from diseased ascending colon tissue removed from a 74-year-old Caucasian male during a multiple-segment large bowel excision with temporary ileostomy. Pathology indicated inflammatory bowel disease consistent with chronic ulcerative colitis, severe acute and chronic mucosal inflammation with erythema, ulceration, and pseudopolyp formation involving the entire colon and rectum. The sigmoid colon had an area of mild stricture formation. One diverticulum with diverticulitis was identified near this zone.
26	PROSBPT03	The library was constructed using RNA isolated from diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy and regional lymph node excision. Pathology indicated benign prostatic hyperplasia (BPH). Pathology for the associated tumor indicated adenocarcinoma, Gleason grade 3+3. The patient presented with elevated prostate specific antigen (PSA), benign hypertension, and hyperlipidemia. Family history included cerebrovascular disease, benign hypertension and prostate cancer.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less, if applicable
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 5 cont.

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13, or a fragment thereof.

2. A substantially purified variant having at least 90% amino acid identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in the sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26, or a fragment thereof.

10. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression of CSIGP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression of CSIGP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
BANDMAN, Olga
HILLMAN, Jennifer L.
LAL, Preeti
YUE, Henry
TANG, Y. Tom
PATTERSON, Chandra
BAUGHN, Mariah R.
YANG, Junming

<120> CELL SIGNALING PROTEINS

<130> PF-0521 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/085,3434; 60/098,010
<151> 1998-05-13; 1998-08-26

<160> 26

<170> PERL Program

<210> 1
<211> 418
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 016108

<400> 1
Met Ser Leu Leu Asp Cys Phe Cys Thr Ser Arg Thr Gln Val Glu
1 5 10 15
Ser Leu Arg Pro Glu Lys Gln Ser Glu Thr Ser Ile His Gln Tyr
20 25 30
Leu Val Asp Glu Pro Thr Leu Ser Trp Ser Arg Pro Ser Thr Arg
35 40 45
Ala Ser Glu Val Leu Cys Ser Thr Asn Val Ser His Tyr Glu Leu
50 55 60
Gln Val Glu Ile Gly Arg Gly Phe Asp Asn Leu Thr Ser Val His
65 70 75
Leu Ala Arg His Thr Pro Thr Gly Thr Leu Val Thr Ile Lys Ile
80 85 90
Thr Asn Leu Glu Asn Cys Asn Glu Glu Arg Leu Lys Ala Leu Gln
95 100 105
Lys Ala Val Ile Leu Ser His Phe Phe Arg His Pro Asn Ile Thr
110 115 120
Thr Tyr Trp Thr Val Phe Thr Val Gly Ser Trp Leu Trp Val Ile
125 130 135
Ser Pro Phe Met Ala Tyr Gly Ser Ala Ser Gln Leu Leu Arg Thr
140 145 150
Tyr Phe Pro Glu Gly Met Ser Glu Thr Leu Ile Arg Asn Ile Leu
155 160 165
Phe Gly Ala Val Arg Gly Leu Asn Tyr Leu His Gln Asn Gly Cys
170 175 180
Ile His Arg Ser Ile Lys Ala Ser His Ile Leu Ile Ser Gly Asp
185 190 195
Gly Leu Val Thr Leu Ser Gly Leu Ser His Leu His Ser Leu Val

```
                    200                     205                     210
Lys His Gly Gln Arg His Arg Ala Val Tyr Asp Phe Pro Gln Phe
                    215                     220                     225
Ser Thr Ser Val Gln Pro Trp Leu Ser Pro Glu Leu Leu Arg Gln
                    230                     235                     240
Asp Leu His Gly Leu Tyr Val Lys Ser Asp Ile Tyr Ser Val Gly
                    245                     250                     255
Ile Thr Ala Cys Glu Leu Ala Ser Gly Gln Val Pro Phe Gln Asp
                    260                     265                     270
Met His Arg Thr Gln Met Leu Leu Gln Lys Leu Lys Gly Pro Pro
                    275                     280                     285
Tyr Ser Pro Leu Asp Ile Ser Ile Phe Pro Gln Ser Glu Ser Arg
                    290                     295                     300
Met Lys Asn Ser Gln Ser Gly Val Asp Ser Gly Ile Gly Glu Ser
                    305                     310                     315
Val Leu Val Ser Ser Gly Thr His Thr Val Asn Ser Asp Arg Leu
                    320                     325                     330
His Thr Pro Ser Ser Lys Thr Phe Ser Pro Ala Phe Phe Ser Leu
                    335                     340                     345
Val Gln Leu Cys Leu Gln Gln Asp Pro Glu Lys Arg Pro Ser Ala
                    350                     355                     360
Ser Ser Leu Leu Ser His Val Phe Phe Lys Gln Met Lys Glu Glu
                    365                     370                     375
Ser Gln Asp Ser Ile Leu Ser Leu Leu Pro Pro Ala Tyr Asn Lys
                    380                     385                     390
Pro Ser Ile Ser Leu Pro Pro Val Leu Pro Trp Thr Glu Pro Glu
                    395                     400                     405
Cys Asp Phe Pro Asp Glu Lys Asp Ser Tyr Trp Glu Phe
                    410                     415
```

<210> 2
<211> 540
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 640521

<400> 2

```
Met Ser Ser Cys Val Ser Ser Gln Pro Ser Ser Asn Arg Ala Ala
 1                  5                   10                  15
Pro Gln Asp Glu Leu Gly Gly Arg Gly Ser Ser Ser Ser Glu Ser
                    20                  25                  30
Gln Lys Pro Cys Glu Ala Leu Arg Gly Leu Ser Ser Leu Ser Ile
                    35                  40                  45
His Leu Gly Met Glu Ser Phe Ile Val Val Thr Glu Cys Glu Pro
                    50                  55                  60
Gly Cys Ala Val Asp Leu Gly Leu Ala Arg Asp Arg Pro Leu Glu
                    65                  70                  75
Ala Asp Gly Gln Glu Val Pro Leu Asp Ser Ser Gly Ser Gln Ala
                    80                  85                  90
Arg Pro His Leu Ser Gly Arg Lys Leu Ser Leu Gln Glu Arg Ser
                    95                  100                 105
Gln Gly Gly Leu Ala Ala Gly Gly Ser Leu Asp Met Asn Gly Arg
                    110                 115                 120
Cys Ile Cys Pro Ser Leu Pro Tyr Ser Pro Val Ser Ser Pro Gln
                    125                 130                 135
Ser Ser Pro Arg Leu Pro Arg Arg Pro Thr Val Glu Ser His His
                    140                 145                 150
Val Ser Ile Thr Gly Met Gln Asp Cys Val Gln Leu Asn Gln Tyr
```

```
                155                     160                     165
Thr Leu Lys Asp Glu Ile Gly Lys Gly Ser Tyr Gly Val Val Lys
                170                     175                     180
Leu Ala Tyr Asn Glu Asn Asp Asn Thr Tyr Tyr Ala Met Lys Val
                185                     190                     195
Leu Ser Lys Lys Lys Leu Ile Arg Gln Ala Gly Phe Pro Arg Arg
                200                     205                     210
Pro Pro Pro Arg Gly Thr Arg Pro Ala Pro Gly Gly Cys Ile Gln
                215                     220                     225
Pro Arg Gly Pro Ile Glu Gln Val Tyr Gln Glu Ile Ala Ile Leu
                230                     235                     240
Lys Lys Leu Asp His Pro Asn Val Val Lys Leu Val Glu Val Leu
                245                     250                     255
Asp Asp Pro Asn Glu Asp His Leu Tyr Met Val Phe Glu Leu Val
                260                     265                     270
Asn Gln Gly Pro Val Met Glu Val Pro Thr Leu Lys Pro Leu Ser
                275                     280                     285
Glu Asp Gln Ala Arg Phe Tyr Phe Gln Asp Leu Ile Lys Gly Ile
                290                     295                     300
Glu Tyr Leu His Tyr Gln Lys Ile Ile His Arg Asp Ile Lys Pro
                305                     310                     315
Ser Asn Leu Leu Val Gly Glu Asp Gly His Ile Lys Ile Ala Asp
                320                     325                     330
Phe Gly Val Ser Asn Glu Phe Lys Gly Ser Asp Ala Leu Leu Ser
                335                     340                     345
Asn Thr Val Gly Thr Pro Ala Phe Met Ala Pro Glu Ser Leu Ser
                350                     355                     360
Glu Thr Arg Lys Ile Phe Ser Gly Lys Ala Leu Asp Val Trp Ala
                365                     370                     375
Met Gly Val Thr Leu Tyr Cys Phe Val Phe Gly Gln Cys Pro Phe
                380                     385                     390
Met Asp Glu Arg Ile Met Cys Leu His Ser Lys Ile Lys Ser Gln
                395                     400                     405
Ala Leu Glu Phe Pro Asp Gln Pro Asp Ile Ala Glu Asp Leu Lys
                410                     415                     420
Asp Leu Ile Thr Arg Met Leu Asp Lys Asn Pro Glu Ser Arg Ile
                425                     430                     435
Val Val Pro Glu Ile Lys Leu His Pro Trp Val Thr Arg His Gly
                440                     445                     450
Ala Glu Pro Leu Pro Ser Glu Asp Glu Asn Cys Thr Leu Val Glu
                455                     460                     465
Val Thr Glu Glu Glu Val Glu Asn Ser Val Lys His Ile Pro Ser
                470                     475                     480
Leu Ala Thr Val Ile Leu Val Lys Thr Met Ile Arg Lys Arg Ser
                485                     490                     495
Phe Gly Asn Pro Phe Glu Gly Ser Arg Arg Glu Glu Arg Ser Leu
                500                     505                     510
Ser Ala Pro Gly Asn Leu Leu Thr Lys Gln Gly Ser Glu Asp Asn
                515                     520                     525
Leu Gln Gly Thr Asp Pro Pro Pro Val Gly Glu Glu Glu Val Leu
                530                     535                     540
```

<210> 3
<211> 729
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1250171

<400> 3
Met Gln Ser Thr Ser Asn His Leu Trp Leu Leu Ser Asp Ile Leu
1   5   10   15
Gly Gln Gly Ala Thr Ala Asn Val Phe Arg Gly Arg His Lys Lys
            20   25   30
Thr Gly Asp Leu Phe Ala Ile Lys Val Phe Asn Asn Ile Ser Phe
            35   40   45
Leu Arg Pro Val Asp Val Gln Met Arg Glu Phe Glu Val Leu Lys
            50   55   60
Lys Leu Asn His Lys Asn Ile Val Lys Leu Phe Ala Ile Glu Glu
            65   70   75
Glu Thr Thr Thr Arg His Lys Val Leu Ile Met Glu Phe Cys Pro
            80   85   90
Cys Gly Ser Leu Tyr Thr Val Leu Glu Glu Pro Ser Asn Ala Tyr
            95   100   105
Gly Leu Pro Glu Ser Glu Phe Leu Ile Val Leu Arg Asp Val Val
            110   115   120
Gly Gly Met Asn His Leu Arg Glu Asn Gly Ile Val His Arg Asp
            125   130   135
Ile Lys Pro Gly Asn Ile Met Arg Val Ile Gly Glu Asp Gly Gln
            140   145   150
Ser Val Tyr Lys Leu Thr Asp Phe Gly Ala Ala Arg Glu Leu Glu
            155   160   165
Asp Asp Glu Gln Phe Val Ser Leu Tyr Gly Thr Glu Glu Tyr Leu
            170   175   180
His Pro Asp Met Tyr Glu Arg Ala Val Leu Arg Lys Asp His Gln
            185   190   195
Lys Lys Tyr Gly Ala Thr Val Asp Leu Trp Ser Ile Gly Val Thr
            200   205   210
Phe Tyr His Ala Ala Thr Gly Ser Leu Pro Phe Arg Pro Phe Glu
            215   220   225
Gly Pro Arg Arg Asn Lys Glu Val Met Tyr Lys Ile Ile Thr Gly
            230   235   240
Lys Pro Ser Gly Ala Ile Ser Gly Val Gln Lys Ala Glu Asn Gly
            245   250   255
Pro Ile Asp Trp Ser Gly Asp Met Pro Val Ser Cys Ser Leu Ser
            260   265   270
Arg Gly Leu Gln Val Leu Leu Thr Pro Val Leu Ala Asn Ile Leu
            275   280   285
Glu Ala Asp Gln Glu Lys Cys Trp Gly Phe Asp Gln Phe Phe Ala
            290   295   300
Glu Thr Ser Asp Ile Leu His Arg Met Val Ile His Val Phe Ser
            305   310   315
Leu Gln Gln Met Thr Ala His Lys Ile Tyr Ile His Ser Tyr Asn
            320   325   330
Thr Ala Thr Ile Phe His Glu Leu Val Tyr Lys Gln Thr Lys Ile
            335   340   345
Ile Ser Ser Asn Gln Glu Leu Ile Tyr Glu Gly Arg Arg Leu Val
            350   355   360
Leu Glu Pro Gly Arg Leu Ala Gln His Phe Pro Lys Thr Thr Glu
            365   370   375
Glu Asn Pro Ile Phe Val Val Ser Arg Glu Pro Leu Asn Thr Ile
            380   385   390
Gly Leu Ile Tyr Glu Lys Ile Ser Leu Pro Lys Val His Pro Arg
            395   400   405
Tyr Asp Leu Asp Gly Asp Ala Ser Met Ala Lys Ala Ile Thr Gly
            410   415   420
Val Val Cys Tyr Ala Cys Arg Ile Ala Ser Thr Leu Leu Leu Tyr
            425   430   435
Gln Glu Leu Met Arg Lys Gly Ile Arg Trp Leu Ile Glu Leu Ile
            440   445   450
Lys Asp Asp Tyr Asn Glu Thr Val His Lys Lys Thr Glu Val Val
            455   460   465
Ile Thr Leu Asp Phe Cys Ile Arg Asn Ile Glu Lys Thr Val Lys

```
                    470                     475                     480
Val Tyr Glu Lys Leu Met Lys Ile Asn Leu Glu Ala Ala Glu Leu
                485                     490                     495
Gly Glu Ile Ser Asp Ile His Thr Lys Leu Leu Arg Leu Ser Ser
                500                     505                     510
Ser Gln Gly Thr Ile Glu Thr Ser Leu Gln Asp Ile Asp Ser Arg
                515                     520                     525
Leu Ser Pro Gly Gly Ser Leu Ala Asp Ala Trp Ala His Gln Glu
                530                     535                     540
Gly Thr His Pro Lys Asp Arg Asn Val Glu Lys Leu Gln Val Leu
                545                     550                     555
Leu Asn Cys Met Thr Glu Ile Tyr Tyr Gln Phe Lys Lys Asp Lys
                560                     565                     570
Ala Glu Arg Arg Leu Ala Tyr Asn Glu Glu Gln Ile His Lys Phe
                575                     580                     585
Asp Lys Gln Lys Leu Tyr Tyr His Ala Thr Lys Ala Met Thr His
                590                     595                     600
Phe Thr Asp Glu Cys Val Lys Lys Tyr Glu Ala Phe Leu Asn Lys
                605                     610                     615
Ser Glu Glu Trp Ile Arg Lys Met Leu His Leu Arg Lys Gln Leu
                620                     625                     630
Leu Ser Leu Thr Asn Gln Cys Phe Asp Ile Glu Glu Glu Val Ser
                635                     640                     645
Lys Tyr Gln Glu Tyr Thr Asn Glu Leu Gln Glu Thr Leu Pro Gln
                650                     655                     660
Lys Met Phe Thr Ala Ser Ser Gly Ile Lys His Thr Met Thr Pro
                665                     670                     675
Ile Tyr Pro Ser Ser Asn Thr Leu Val Glu Met Thr Leu Gly Met
                680                     685                     690
Lys Lys Leu Lys Glu Glu Met Glu Gly Val Val Lys Glu Leu Ala
                695                     700                     705
Glu Asn Asn His Ile Leu Glu Arg Phe Gly Ser Leu Thr Met Asp
                710                     715                     720
Gly Gly Leu Arg Asn Val Asp Cys Leu
                725
```

```
<210> 4
<211> 313
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1911587

<400> 4
Met Pro Gly Leu Leu Leu Cys Glu Pro Thr Glu Leu Tyr Asn Ile
  1               5                  10                  15
Leu Asn Gln Ala Thr Lys Leu Ser Arg Leu Thr Asp Pro Asn Tyr
                 20                  25                  30
Leu Cys Leu Leu Asp Val Arg Ser Lys Trp Glu Tyr Asp Glu Ser
                 35                  40                  45
His Val Ile Thr Ala Leu Arg Val Lys Lys Lys Asn Asn Glu Tyr
                 50                  55                  60
Leu Leu Pro Glu Ser Val Asp Leu Glu Cys Val Lys Tyr Cys Val
                 65                  70                  75
Val Tyr Asp Asn Asn Ser Ser Thr Leu Glu Ile Leu Leu Lys Asp
                 80                  85                  90
Asp Asp Asp Asp Ser Asp Ser Asp Gly Asp Gly Lys Asp Leu Val
                 95                 100                 105
Pro Gln Ala Ala Ile Glu Tyr Gly Arg Ile Leu Thr Arg Leu Thr
```

110 115 120
His His Pro Val Tyr Ile Leu Lys Gly Gly Tyr Glu Arg Phe Ser
125 130 135
Gly Thr Tyr His Phe Leu Arg Thr Gln Lys Ile Ile Trp Met Pro
140 145 150
Gln Glu Leu Asp Ala Phe Gln Pro Tyr Pro Ile Glu Ile Val Pro
155 160 165
Gly Lys Val Phe Val Gly Asn Phe Ser Gln Ala Cys Asp Pro Lys
170 175 180
Ile Gln Lys Asp Leu Lys Ile Lys Ala His Val Asn Val Ser Met
185 190 195
Asp Thr Gly Pro Phe Phe Ala Gly Asp Ala Asp Arg Leu Leu His
200 205 210
Ile Arg Ile Glu Asp Ser Pro Glu Ala Gln Ile Leu Pro Phe Leu
215 220 225
Arg His Met Cys His Phe Ile Glu Ile His His His Leu Gly Ser
230 235 240
Val Ile Leu Ile Phe Ser Thr Gln Gly Ile Ser Arg Ser Cys Ala
245 250 255
Ala Ile Ile Ala Tyr Leu Met His Ser Asn Glu Gln Thr Leu Gln
260 265 270
Arg Ser Trp Ala Tyr Val Lys Lys Cys Lys Asn Asn Met Cys Pro
275 280 285
Asn Arg Gly Leu Val Ser Gln Leu Leu Glu Trp Glu Lys Thr Ile
290 295 300
Leu Gly Asp Ser Ile Thr Asn Ile Met Asp Pro Leu Tyr
305 310

<210> 5
<211> 506
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 2079081

<400> 5
Met Arg Asp Pro Leu Thr Asp Cys Pro Tyr Asn Lys Val Tyr Lys
1 5 10 15
Asn Leu Lys Glu Phe Ser Gln Asn Gly Glu Asn Phe Cys Lys Gln
20 25 30
Val Thr Ser Val Leu Gln Gln Arg Ala Asn Leu Glu Ile Ser Tyr
35 40 45
Ala Lys Gly Leu Gln Lys Leu Ala Ser Lys Leu Ser Lys Ala Leu
50 55 60
Gln Asn Thr Arg Lys Ser Cys Val Ser Ser Ala Trp Ala Trp Ala
65 70 75
Ser Glu Gly Met Lys Ser Thr Ala Asp Leu His Gln Lys Leu Gly
80 85 90
Lys Ala Ile Glu Leu Glu Ala Ile Lys Pro Thr Tyr Gln Val Leu
95 100 105
Asn Val Gln Glu Lys Lys Arg Lys Ser Leu Asp Asn Glu Val Glu
110 115 120
Lys Thr Ala Asn Leu Val Ile Ser Asn Trp Asn Gln Gln Ile Lys
125 130 135
Ala Lys Lys Lys Leu Met Val Ser Thr Lys Lys His Glu Ala Leu
140 145 150
Phe Gln Leu Val Glu Ser Ser Lys Gln Ser Met Thr Glu Lys Glu
155 160 165
Lys Arg Lys Leu Leu Asn Lys Leu Thr Lys Ser Thr Glu Lys Leu

170 175 180
Glu Lys Glu Asp Glu Asn Tyr Tyr Gln Lys Asn Met Ala Gly Tyr
185 190 195
Ser Thr Arg Leu Lys Trp Glu Asn Thr Leu Glu Asn Cys Tyr Gln
200 205 210
Ser Ile Leu Glu Leu Glu Lys Glu Arg Ile Gln Leu Leu Cys Asn
215 220 225
Asn Leu Asn Gln Tyr Ser Gln His Ile Ser Leu Phe Gly Gln Thr
230 235 240
Leu Thr Thr Cys His Thr Gln Ile His Cys Ala Ile Ser Lys Ile
245 250 255
Asp Ile Glu Lys Asp Ile Gln Ala Val Met Glu Glu Thr Ala Ile
260 265 270
Leu Ser Thr Glu Asn Lys Ser Glu Phe Leu Leu Thr Asp Tyr Phe
275 280 285
Glu Glu Asp Pro Asn Ser Ala Met Asp Lys Glu Arg Arg Lys Ser
290 295 300
Leu Leu Lys Pro Lys Leu Leu Arg Leu Gln Arg Asp Ile Glu Lys
305 310 315
Ala Ser Lys Asp Lys Glu Gly Leu Glu Arg Met Leu Lys Thr Tyr
320 325 330
Ser Ser Thr Ser Ser Phe Ser Asp Ala Lys Ser Gln Lys Asp Thr
335 340 345
Ala Ala Leu Met Asp Glu Asn Asn Leu Lys Leu Asp Leu Leu Glu
350 355 360
Ala Asn Ser Tyr Lys Leu Ser Ser Met Leu Ala Glu Leu Glu Gln
365 370 375
Arg Pro Gln Pro Ser His Pro Cys Ser Asn Ser Ile Phe Arg Trp
380 385 390
Arg Glu Lys Glu His Thr His Ser Tyr Val Lys Ile Ser Arg Pro
395 400 405
Phe Leu Met Lys Arg Leu Glu Asn Ile Val Ser Lys Ala Ser Ser
410 415 420
Gly Gly Gln Ser Asn Pro Gly Ser Ser Thr Pro Ala Pro Gly Ala
425 430 435
Ala Gln Leu Ser Ser Arg Leu Cys Lys Ala Leu Tyr Ser Phe Gln
440 445 450
Ala Arg Gln Asp Asp Glu Leu Asn Leu Glu Lys Gly Asp Ile Val
455 460 465
Ile Ile His Glu Lys Lys Glu Glu Gly Trp Trp Phe Gly Ser Leu
470 475 480
Asn Gly Lys Lys Gly His Phe Pro Ala Ala Tyr Val Glu Glu Leu
485 490 495
Pro Ser Asn Ala Gly Asn Thr Ala Thr Lys Ala
500 505

<210> 6
<211> 341
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 2472655

<400> 6
Met Arg Lys Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly
1 5 10 15
Leu Ala Leu Cys Lys Arg Leu Leu Ala Glu Asp Asp Glu Leu His
20 25 30
Leu Cys Leu Ala Cys Arg Asn Met Ser Lys Ala Glu Ala Val Cys

```
                    35                  40                  45
Ala Ala Leu Leu Ala Ser His Pro Thr Ala Glu Val Thr Ile Val
                    50                  55                  60
Gln Val Asp Val Ser Asn Leu Gln Ser Val Phe Arg Ala Ser Lys
                    65                  70                  75
Glu Leu Lys Gln Arg Phe Gln Arg Leu Asp Cys Ile Tyr Leu Asn
                    80                  85                  90
Ala Gly Ile Met Pro Asn Pro Gln Leu Asn Ile Lys Ala Leu Phe
                    95                  100                 105
Phe Gly Leu Phe Ser Arg Lys Val Ile His Met Phe Ser Thr Ala
                    110                 115                 120
Glu Gly Leu Leu Thr Gln Gly Asp Lys Ile Thr Ala Asp Gly Leu
                    125                 130                 135
Gln Glu Val Phe Glu Thr Asn Val Phe Gly His Phe Ile Leu Ile
                    140                 145                 150
Arg Glu Leu Glu Pro Leu Leu Cys His Ser Asp Asn Pro Ser Gln
                    155                 160                 165
Leu Ile Trp Thr Ser Ser Arg Ser Ala Arg Lys Ser Asn Phe Ser
                    170                 175                 180
Leu Glu Asp Phe Gln His Ser Lys Gly Lys Glu Pro Tyr Ser Ser
                    185                 190                 195
Ser Lys Tyr Ala Thr Asp Leu Leu Ser Val Ala Leu Asn Arg Asn
                    200                 205                 210
Phe Asn Gln Gln Gly Leu Tyr Ser Asn Val Ala Cys Pro Gly Thr
                    215                 220                 225
Ala Leu Thr Asn Leu Thr Tyr Gly Ile Leu Pro Pro Phe Ile Trp
                    230                 235                 240
Thr Leu Leu Met Pro Ala Ile Leu Leu Leu Arg Phe Phe Ala Asn
                    245                 250                 255
Ala Phe Thr Leu Thr Pro Tyr Asn Gly Thr Glu Ala Leu Val Trp
                    260                 265                 270
Leu Phe His Gln Lys Pro Glu Ser Leu Asn Pro Leu Ile Lys Tyr
                    275                 280                 285
Leu Ser Ala Thr Thr Gly Phe Gly Arg Asn Tyr Ile Met Thr Gln
                    290                 295                 300
Lys Met Asp Leu Asp Glu Asp Thr Ala Glu Lys Phe Tyr Gln Lys
                    305                 310                 315
Leu Leu Glu Leu Glu Lys His Ile Arg Val Thr Ile Gln Lys Thr
                    320                 325                 330
Asp Asn Gln Ala Arg Leu Ser Gly Ser Cys Leu
                    335                 340
```

<210> 7
<211> 898
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 2948818

<400> 7

```
Met Arg Lys Gly Val Leu Lys Asp Pro Glu Ile Ala Asp Leu Ser
 1               5                  10                  15
Tyr Lys Asp Asp Pro Glu Glu Leu Phe Ile Gly Leu His Glu Ile
                20                  25                  30
Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Thr Asn Ala His
                35                  40                  45
Thr Ser Glu Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys
                50                  55                  60
Gln Thr His Glu Lys Trp Gln Asp Ile Leu Lys Glu Val Lys Phe
```

65 70 75
Leu Arg Gln Leu Lys His Pro Asn Thr Ile Glu Tyr Lys Gly Cys
80 85 90
Tyr Leu Lys Glu His Thr Ala Trp Leu Val Met Glu Tyr Cys Leu
95 100 105
Gly Ser Ala Ser Asp Leu Leu Glu Val His Lys Lys Pro Leu Gln
110 115 120
Glu Val Glu Ile Ala Ala Ile Thr His Gly Ala Leu His Gly Leu
125 130 135
Ala Tyr Leu His Ser His Ala Leu Ile His Arg Asp Ile Lys Ala
140 145 150
Gly Asn Ile Leu Leu Thr Glu Pro Gly Gln Val Lys Leu Ala Asp
155 160 165
Phe Gly Ser Ala Ser Met Ala Ser Pro Ala Asn Ser Phe Val Gly
170 175 180
Thr Pro Tyr Trp Met Ala Pro Glu Val Ile Leu Ala Met Asp Glu
185 190 195
Gly Gln Tyr Asp Gly Lys Val Asp Ile Trp Ser Leu Gly Ile Thr
200 205 210
Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro Leu Phe Asn Met Asn
215 220 225
Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn Asp Ser Pro Thr
230 235 240
Leu Gln Ser Asn Glu Trp Thr Asp Ser Phe Arg Arg Phe Val Asp
245 250 255
Tyr Cys Leu Gln Lys Ile Pro Gln Glu Arg Pro Thr Ser Ala Glu
260 265 270
Leu Leu Arg His Asp Phe Val Arg Arg Asp Arg Pro Leu Arg Val
275 280 285
Leu Ile Asp Leu Ile Gln Arg Thr Lys Asp Ala Val Arg Glu Leu
290 295 300
Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Ile Leu Phe Gln Glu
305 310 315
Thr Arg Asn Gly Pro Leu Asn Glu Ser Gln Glu Asp Glu Glu Asp
320 325 330
Ser Glu His Gly Thr Ser Leu Asn Arg Glu Met Asp Ser Leu Gly
335 340 345
Ser Asn His Ser Ile Pro Ser Met Ser Val Ser Thr Gly Ser Gln
350 355 360
Ser Ser Ser Val Asn Ser Met Gln Glu Val Met Asp Glu Ser Ser
365 370 375
Ser Glu Leu Val Met Met His Asp Asp Glu Ser Thr Ile Asn Ser
380 385 390
Ser Ser Ser Val Val His Lys Lys Asp His Val Phe Ile Arg Asp
395 400 405
Glu Ala Gly His Gly Asp Pro Arg Pro Glu Pro Arg Pro Thr Gln
410 415 420
Ser Val Gln Ser Gln Ala Leu His Tyr Arg Asn Arg Glu Arg Phe
425 430 435
Ala Thr Ile Lys Ser Ala Ser Leu Val Thr Arg Gln Ile His Glu
440 445 450
His Glu Gln Glu Asn Glu Leu Arg Glu Gln Met Ser Gly Tyr Lys
455 460 465
Arg Met Arg Arg Gln His Gln Lys Gln Leu Ile Ala Leu Glu Asn
470 475 480
Lys Leu Lys Ala Glu Met Asp Glu His Arg Leu Lys Leu Gln Lys
485 490 495
Glu Val Glu Thr His Ala Asn Asn Ser Ser Ile Glu Leu Glu Lys
500 505 510
Leu Ala Lys Lys Gln Val Ala Ile Ile Glu Lys Glu Ala Lys Val
515 520 525
Ala Ala Ala Asp Glu Lys Lys Phe Gln Gln Gln Ile Leu Ala Gln
530 535 540
Gln Lys Lys Asp Leu Thr Thr Phe Leu Glu Ser Gln Lys Lys Gln

545 550 555
Tyr Lys Ile Cys Lys Glu Lys Ile Lys Glu Glu Met Asn Glu Asp
560 565 570
His Ser Thr Pro Lys Lys Glu Lys Gln Glu Arg Ile Ser Lys His
575 580 585
Lys Glu Asn Leu Gln His Thr Gln Ala Glu Glu Glu Ala His Leu
590 595 600
Leu Thr Gln Gln Arg Leu Tyr Tyr Asp Lys Asn Cys Arg Phe Phe
605 610 615
Lys Arg Lys Ile Met Ile Lys Arg His Glu Val Glu Gln Gln Asn
620 625 630
Ile Arg Glu Glu Leu Asn Lys Lys Arg Thr Gln Lys Glu Met Glu
635 640 645
His Ala Met Leu Ile Arg His Asp Glu Ser Thr Arg Glu Leu Glu
650 655 660
Tyr Arg Gln Leu His Thr Leu Gln Lys Leu Arg Met Asp Leu Ile
665 670 675
Arg Leu Gln His Gln Thr Glu Leu Glu Asn Gln Leu Glu Tyr Asn
680 685 690
Lys Arg Arg Glu Arg Glu Leu His Arg Lys His Val Met Glu Leu
695 700 705
Arg Gln Gln Pro Lys Asn Leu Lys Ala Met Glu Met Gln Ile Lys
710 715 720
Lys Gln Phe Gln Asp Thr Cys Lys Val Gln Thr Lys Gln Tyr Lys
725 730 735
Ala Leu Lys Asn His Gln Leu Glu Val Thr Pro Lys Asn Glu His
740 745 750
Lys Thr Ile Leu Lys Thr Leu Lys Asp Glu Gln Thr Arg Lys Leu
755 760 765
Ala Ile Leu Ala Glu Gln Tyr Glu Gln Ser Ile Asn Glu Met Met
770 775 780
Ala Ser Gln Ala Leu Arg Leu Asp Glu Ala Gln Glu Ala Glu Cys
785 790 795
Gln Ala Leu Arg Leu Gln Leu Gln Gln Glu Met Glu Leu Leu Asn
800 805 810
Ala Tyr Gln Ser Lys Ile Lys Met Gln Thr Glu Ala Gln His Glu
815 820 825
Arg Glu Leu Gln Lys Leu Glu Gln Arg Val Ser Leu Arg Arg Ala
830 835 840
His Leu Glu Gln Lys Ile Glu Glu Glu Leu Ala Ala Leu Gln Lys
845 850 855
Glu Arg Ser Glu Arg Ile Lys Asn Leu Leu Glu Arg Gln Glu Arg
860 865 870
Glu Ile Glu Thr Phe Asp Met Glu Ser Leu Arg Met Gly Phe Gly
875 880 885
Asn Leu Val Thr Leu Asp Phe Pro Lys Glu Asp Tyr Arg
890 895

<210> 8
<211> 336
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 054191

<400> 8
Met Ala Thr Leu Ser Val Ile Gly Ser Ser Ser Leu Ile Ala Tyr
1 5 10 15
Ala Val Phe His Asn Ile Gln Lys Ser Pro Glu Ile Arg Pro Leu

```
                    20                          25                          30
Phe Tyr Leu Ser Phe Cys Asp Leu Leu Leu Gly Leu Cys Trp Leu
                    35                          40                          45
Thr Glu Thr Leu Leu Tyr Gly Ala Ser Val Ala Asn Lys Asp Ile
                    50                          55                          60
Ile Cys Tyr Asn Leu Gln Ala Val Gly Gln Ile Phe Tyr Ile Ser
                    65                          70                          75
Ser Phe Leu Tyr Thr Val Asn Tyr Ile Trp Tyr Leu Tyr Thr Glu
                    80                          85                          90
Leu Arg Met Lys His Thr Gln Ser Gly Gln Ser Thr Ser Pro Leu
                    95                          100                         105
Val Ile Asp Tyr Thr Cys Arg Val Gly Gln Met Ala Phe Val Phe
                    110                         115                         120
Ser Ser Leu Ile Pro Leu Leu Leu Met Thr Pro Val Phe Cys Leu
                    125                         130                         135
Gly Asn Thr Ser Glu Cys Phe Gln Asn Phe Ser Gln Ser His Lys
                    140                         145                         150
Cys Ile Leu Met His Ser Pro Pro Ser Ala Met Ala Glu Leu Pro
                    155                         160                         165
Pro Ser Ala Asn Thr Ser Val Cys Ser Thr Leu Tyr Phe Tyr Gly
                    170                         175                         180
Ile Ala Ile Phe Leu Gly Ser Phe Val Leu Ser Leu Leu Thr Ile
                    185                         190                         195
Met Val Leu Leu Ile Arg Ala Gln Thr Leu Tyr Lys Lys Phe Val
                    200                         205                         210
Lys Ser Thr Gly Phe Leu Gly Ser Glu Gln Trp Ala Val Ile His
                    215                         220                         225
Ile Val Asp Gln Arg Val Arg Phe Tyr Pro Val Ala Phe Phe Cys
                    230                         235                         240
Cys Trp Gly Pro Ala Val Ile Leu Met Ile Ile Lys Leu Thr Lys
                    245                         250                         255
Pro Gln Asp Thr Lys Leu His Met Ala Leu Tyr Val Leu Gln Ala
                    260                         265                         270
Leu Thr Ala Thr Ser Gln Gly Leu Leu Asn Cys Gly Val Tyr Gly
                    275                         280                         285
Trp Thr Gln His Lys Phe His Gln Leu Lys Gln Glu Ala Arg Arg
                    290                         295                         300
Asp Ala Asp Thr Gln Thr Pro Leu Leu Cys Ser Gln Lys Arg Phe
                    305                         310                         315
Tyr Ser Arg Gly Leu Asn Ser Leu Glu Ser Thr Leu Thr Phe Pro
                    320                         325                         330
Ala Ser Thr Ser Thr Ile
                    335
```

```
<210> 9
<211> 686
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1403604

<400> 9
Met Gly Pro Arg Ser Arg Glu Arg Arg Ala Gly Ala Val Gln Asn
  1                 5                           10                          15
Thr Asn Asp Ser Ser Ala Leu Ser Lys Arg Ser Leu Ala Ala Arg
                    20                          25                          30
Gly Tyr Val Gln Asp Pro Phe Ala Ala Leu Leu Val Pro Gly Ala
                    35                          40                          45
Ala Arg Arg Ala Pro Leu Ile His Arg Gly Tyr Tyr Val Arg Ala
```

```
                    50                  55                  60
Arg Ala Val Arg His Cys Val Arg Ala Phe Leu Glu Gln Ile Gly
                65                  70                  75
Ala Pro Gln Ala Ala Leu Arg Ala Gln Ile Leu Ser Leu Gly Ala
                80                  85                  90
Gly Phe Asp Ser Leu Tyr Phe Arg Leu Lys Thr Ala Gly Arg Leu
                95                  100                 105
Ala Arg Ala Ala Val Trp Glu Val Asp Phe Pro Asp Val Ala Arg
                110                 115                 120
Arg Lys Ala Glu Arg Ile Gly Glu Thr Pro Glu Leu Cys Ala Leu
                125                 130                 135
Thr Gly Pro Phe Glu Arg Gly Glu Pro Ala Ser Ala Leu Cys Phe
                140                 145                 150
Glu Ser Ala Asp Tyr Cys Ile Leu Gly Leu Asp Leu Arg Gln Leu
                155                 160                 165
Gln Arg Val Glu Glu Ala Leu Gly Ala Ala Gly Leu Asp Ala Ala
                170                 175                 180
Ser Pro Thr Leu Leu Leu Ala Glu Ala Val Leu Thr Tyr Leu Glu
                185                 190                 195
Pro Glu Ser Ala Ala Ala Leu Ile Ala Trp Ala Ala Gln Arg Phe
                200                 205                 210
Pro Asn Ala Leu Phe Val Val Tyr Glu Gln Met Arg Pro Gln Asp
                215                 220                 225
Ala Phe Gly Gln Phe Met Leu Gln His Phe Arg Gln Leu Asn Ser
                230                 235                 240
Pro Leu His Gly Leu Glu Arg Phe Pro Asp Val Glu Ala Gln Arg
                245                 250                 255
Arg Arg Phe Leu Gln Ala Gly Trp Thr Ala Cys Gly Ala Val Asp
                260                 265                 270
Ile Asn Glu Phe Tyr His Cys Phe Leu Pro Ala Glu Glu Arg Arg
                275                 280                 285
Arg Val Glu Asn Ile Glu Pro Phe Asp Glu Phe Glu Glu Trp His
                290                 295                 300
Leu Lys Cys Ala His Tyr Phe Ile Leu Ala Ala Ser Arg Gly Asp
                305                 310                 315
Thr Leu Ser His Thr Leu Val Phe Pro Ser Ser Glu Ala Phe Pro
                320                 325                 330
Arg Val Asn Pro Ala Ser Pro Ser Gly Val Phe Pro Ala Ser Val
                335                 340                 345
Val Ser Ser Glu Gly Gln Val Pro Asn Leu Lys Arg Tyr Gly His
                350                 355                 360
Ala Ser Val Phe Leu Ser Pro Asp Val Ile Leu Ser Ala Gly Gly
                365                 370                 375
Phe Gly Glu Gln Glu Gly Arg His Cys Arg Val Ser Gln Phe His
                380                 385                 390
Leu Leu Ser Arg Asp Cys Asp Ser Glu Trp Lys Gly Ser Gln Ile
                395                 400                 405
Gly Ser Cys Gly Thr Gly Val Gln Trp Asp Gly Arg Leu Tyr His
                410                 415                 420
Thr Met Thr Arg Leu Ser Glu Ser Arg Val Leu Val Leu Gly Gly
                425                 430                 435
Arg Leu Ser Pro Val Ser Pro Ala Leu Gly Val Leu Gln Leu His
                440                 445                 450
Phe Phe Lys Ser Glu Asp Asn Asn Thr Glu Asp Leu Lys Val Thr
                455                 460                 465
Ile Thr Lys Ala Gly Arg Lys Asp Asp Ser Thr Leu Cys Cys Trp
                470                 475                 480
Arg His Ser Thr Thr Glu Val Ser Cys Gln Asn Gln Glu Tyr Leu
                485                 490                 495
Phe Val Tyr Gly Gly Arg Ser Val Val Glu Pro Val Leu Ser Asp
                500                 505                 510
Trp His Phe Leu His Val Gly Thr Met Ala Trp Val Arg Ile Pro
                515                 520                 525
Val Glu Gly Glu Val Pro Glu Ala Arg His Ser His Ser Ala Cys
```

```
                    530                     535                     540
Thr Trp Gln Gly Gly Ala Leu Ile Ala Gly Gly Leu Gly Ala Ser
                545                     550                     555
Glu Glu Pro Leu Asn Ser Val Leu Phe Leu Arg Pro Ile Ser Cys
                560                     565                     570
Gly Phe Leu Trp Glu Ser Val Asp Ile Gln Pro Pro Ile Thr Pro
                575                     580                     585
Arg Tyr Ser His Thr Ala His Val Leu Asn Gly Lys Leu Leu Leu
                590                     595                     600
Val Gly Gly Ile Trp Ile His Ser Ser Ser Phe Pro Gly Val Thr
                605                     610                     615
Val Ile Asn Leu Thr Thr Gly Leu Ser Ser Glu Tyr Gln Ile Asp
                620                     625                     630
Thr Thr Tyr Val Pro Trp Pro Leu Met Leu His Asn His Thr Ser
                635                     640                     645
Ile Leu Leu Pro Glu Glu Gln Gln Leu Leu Leu Leu Gly Gly Gly
                650                     655                     660
Gly Asn Cys Phe Ser Phe Gly Thr Tyr Phe Asn Pro His Thr Val
                665                     670                     675
Thr Leu Asp Leu Ser Ser Leu Ser Ala Gly Gln
                680                     685
```

<210> 10
<211> 519
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1652936

<400> 10

```
Met Met Ser Lys Asn Asp Gly Glu Ile Arg Phe Gly Asn Pro Ala
 1               5                  10                  15
Glu Leu His Gly Thr Lys Val Gln Ile Pro Tyr Leu Thr Thr Glu
                20                  25                  30
Lys Asn Ser Phe Lys Arg Met Asp Asp Glu Asp Lys Gln Glu Thr
                35                  40                  45
Gln Ser Pro Thr Met Ser Pro Leu Ala Ser Pro Pro Ser Ser Pro
                50                  55                  60
Pro His Tyr Gln Arg Val Pro Leu Ser His Gly Tyr Ser Lys Leu
                65                  70                  75
Arg Ser Ser Ala Glu Gln Met His Pro Ala Pro Tyr Glu Ala Arg
                80                  85                  90
Gln Pro Leu Val Gln Pro Glu Gly Ser Ser Ser Gly Gly Pro Gly
                95                 100                 105
Thr Lys Pro Leu Arg His Gln Ala Ser Leu Ile Arg Ser Phe Ser
               110                 115                 120
Val Glu Arg Glu Leu Gln Asp Asn Ser Ser Tyr Pro Asp Glu Pro
               125                 130                 135
Trp Arg Ile Thr Glu Glu Gln Arg Glu Tyr Tyr Val Asn Gln Phe
               140                 145                 150
Arg Ser Leu Gln Pro Asp Pro Ser Ser Phe Ile Ser Gly Ser Val
               155                 160                 165
Ala Lys Asn Phe Phe Thr Lys Ser Lys Leu Ser Ile Pro Glu Leu
               170                 175                 180
Ser Tyr Ile Trp Glu Leu Ser Asp Ala Asp Cys Asp Gly Ala Leu
               185                 190                 195
Thr Leu Pro Glu Phe Cys Ala Ala Phe His Leu Ile Val Ala Arg
               200                 205                 210
Lys Asn Gly Tyr Pro Leu Pro Glu Gly Leu Pro Pro Thr Leu Gln
```

215 220 225
Pro Glu Tyr Leu Gln Ala Ala Phe Pro Lys Pro Lys Trp Asp Cys
230 235 240
Gln Leu Phe Asp Ser Tyr Ser Glu Ser Leu Pro Ala Asn Gln Gln
245 250 255
Pro Arg Asp Leu Asn Arg Met Glu Thr Ser Val Lys Asp Met Ala
260 265 270
Asp Leu Pro Val Pro Asn Gln Asp Val Thr Ser Asp Asp Lys Gln
275 280 285
Ala Leu Lys Ser Thr Ile Asn Glu Ala Leu Pro Lys Asp Val Ser
290 295 300
Glu Asp Pro Ala Thr Pro Lys Asp Ser Asn Ser Leu Lys Ala Arg
305 310 315
Pro Arg Ser Arg Ser Tyr Ser Ser Thr Ser Ile Glu Glu Ala Met
320 325 330
Lys Arg Gly Glu Asp Pro Pro Thr Pro Pro Pro Arg Pro Gln Lys
335 340 345
Thr His Ser Arg Ala Ser Ser Leu Asp Leu Asn Lys Val Phe Gln
350 355 360
Pro Ser Val Pro Ala Thr Lys Ser Gly Leu Leu Pro Pro Pro Pro
365 370 375
Ala Leu Pro Pro Arg Pro Cys Pro Ser Gln Ser Glu Gln Val Ser
380 385 390
Glu Ala Glu Leu Leu Pro Gln Leu Ser Arg Ala Pro Ser Gln Ala
395 400 405
Ala Glu Ser Ser Pro Ala Lys Lys Asp Val Leu Tyr Ser Gln Pro
410 415 420
Pro Ser Lys Pro Ile Arg Arg Lys Phe Arg Pro Glu Asn Gln Ala
425 430 435
Thr Glu Asn Gln Glu Pro Ser Thr Ala Ala Ser Gly Pro Ala Ser
440 445 450
Ala Ala Thr Met Lys Pro His Pro Thr Val Gln Lys Gln Ser Ser
455 460 465
Lys Gln Lys Lys Ala Ile Gln Thr Ala Ile Arg Lys Asn Lys Glu
470 475 480
Ala Asn Ala Val Leu Ala Arg Leu Asn Ser Glu Leu Gln Gln Gln
485 490 495
Leu Lys Glu Val His Gln Glu Arg Ile Ala Leu Glu Asn Gln Leu
500 505 510
Glu Gln Leu Arg Pro Val Thr Val Leu
515

<210> 11
<211> 334
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1710702

<400> 11
Met Phe Arg Trp Glu Arg Ser Ile Pro Leu Arg Gly Ser Ala Ala
1 5 10 15
Ala Leu Cys Asn Asn Leu Ser Val Leu Gln Leu Pro Ala Arg Asn
20 25 30
Leu Thr Tyr Phe Gly Val Val His Gly Pro Ser Ala Gln Leu Leu
35 40 45
Ser Ala Ala Pro Glu Gly Val Pro Leu Ala Gln Arg Gln Leu His
50 55 60
Ala Lys Glu Gly Ala Gly Val Ser Pro Pro Leu Ile Thr Gln Val

```
                    65                      70                      75
His Trp Cys Val Leu Pro Phe Arg Val Leu Leu Val Leu Thr Ser
                    80                      85                      90
His Arg Gly Ile Gln Met Tyr Glu Ser Asn Gly Tyr Thr Met Val
                    95                      100                     105
Tyr Trp His Ala Leu Asp Ser Gly Asp Ala Ser Pro Val Gln Ala
                    110                     115                     120
Val Phe Ala Arg Gly Ile Ala Ala Ser Gly His Phe Ile Cys Val
                    125                     130                     135
Gly Thr Trp Ser Gly Arg Val Leu Val Phe Asp Ile Pro Ala Lys
                    140                     145                     150
Gly Pro Asn Ile Val Leu Ser Glu Glu Leu Ala Gly His Gln Met
                    155                     160                     165
Pro Ile Thr Asp Ile Ala Thr Glu Pro Ala Gln Gly Gln Asp Cys
                    170                     175                     180
Val Ala Asp Met Val Thr Ala Asp Asp Ser Gly Leu Leu Cys Val
                    185                     190                     195
Trp Arg Ser Gly Pro Glu Phe Thr Leu Leu Thr Arg Ile Pro Gly
                    200                     205                     210
Phe Gly Val Pro Cys Pro Ser Val Gln Leu Trp Gln Gly Ile Ile
                    215                     220                     225
Ala Ala Gly Tyr Gly Asn Gly Gln Val His Leu Tyr Glu Ala Thr
                    230                     235                     240
Thr Gly Asn Leu His Val Gln Ile Asn Ala His Ala Arg Ala Ile
                    245                     250                     255
Cys Ala Leu Asp Leu Ala Ser Glu Val Gly Lys Leu Leu Ser Ala
                    260                     265                     270
Gly Glu Asp Thr Phe Val His Ile Trp Lys Leu Ser Arg Asn Pro
                    275                     280                     285
Glu Ser Gly Tyr Ile Glu Val Glu His Cys His Gly Glu Cys Val
                    290                     295                     300
Ala Asp Thr Gln Leu Cys Gly Ala Arg Phe Cys Asp Ser Ser Gly
                    305                     310                     315
Asn Ser Phe Ala Val Thr Gly Tyr Asp Leu Ala Glu Ile Arg Arg
                    320                     325                     330
Phe Ser Ser Val
```

<210> 12
<211> 569
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 3239149

<400> 12

```
Met Asp Pro Ala Glu Ala Val Leu Gln Glu Lys Ala Leu Lys Phe
 1                  5                       10                      15
Met Asn Ser Ser Glu Arg Glu Asp Cys Asn Asn Gly Glu Pro Pro
                    20                      25                      30
Arg Lys Ile Ile Pro Glu Lys Asn Ser Leu Arg Gln Thr Tyr Asn
                    35                      40                      45
Ser Cys Ala Arg Leu Cys Leu Asn Gln Glu Thr Val Cys Leu Ala
                    50                      55                      60
Ser Thr Ala Met Lys Thr Glu Asn Cys Val Ala Lys Thr Lys Leu
                    65                      70                      75
Ala Asn Gly Thr Ser Ser Met Ile Val Pro Lys Gln Arg Lys Leu
                    80                      85                      90
Ser Ala Ser Tyr Glu Lys Glu Lys Glu Leu Cys Val Lys Tyr Phe
                    95                      100                     105
```

Glu Gln Trp Ser Glu Ser Asp Gln Val Glu Phe Val Glu His Leu
110 115 120
Ile Ser Gln Met Cys His Tyr Gln His Gly His Ile Asn Ser Tyr
125 130 135
Leu Lys Pro Met Leu Gln Arg Asp Phe Ile Thr Ala Leu Pro Ala
140 145 150
Arg Gly Leu Asp His Ile Ala Glu Asn Ile Leu Ser Tyr Leu Asp
155 160 165
Ala Lys Ser Leu Cys Ala Ala Glu Leu Val Cys Lys Glu Trp Tyr
170 175 180
Arg Val Thr Ser Asp Gly Met Leu Trp Lys Lys Leu Ile Glu Arg
185 190 195
Met Val Arg Thr Asp Ser Leu Trp Arg Gly Leu Ala Glu Arg Arg
200 205 210
Gly Trp Gly Gln Tyr Leu Phe Lys Asn Lys Pro Pro Asp Gly Asn
215 220 225
Ala Pro Pro Asn Ser Phe Tyr Arg Ala Leu Tyr Pro Lys Ile Ile
230 235 240
Gln Asp Ile Glu Thr Ile Glu Ser Asn Trp Arg Cys Gly Arg His
245 250 255
Ser Leu Gln Arg Ile His Cys Arg Ser Glu Thr Ser Lys Gly Val
260 265 270
Tyr Cys Leu Gln Tyr Asp Asp Gln Lys Ile Val Ser Gly Leu Arg
275 280 285
Asp Asn Thr Ile Lys Ile Trp Asp Lys Asn Thr Leu Glu Cys Lys
290 295 300
Arg Ile Leu Thr Gly His Thr Gly Ser Val Leu Cys Leu Gln Tyr
305 310 315
Asp Glu Arg Val Ile Ile Thr Gly Ser Ser Asp Ser Thr Val Arg
320 325 330
Val Trp Asp Val Asn Thr Gly Glu Met Leu Asn Thr Leu Ile His
335 340 345
His Cys Glu Ala Val Leu His Leu Arg Phe Asn Asn Gly Met Met
350 355 360
Val Thr Cys Ser Lys Asp Arg Ser Ile Ala Val Trp Asp Met Ala
365 370 375
Ser Pro Thr Asp Ile Thr Leu Arg Arg Val Leu Val Gly His Arg
380 385 390
Ala Ala Val Asn Val Val Asp Phe Asp Asp Lys Tyr Ile Val Ser
395 400 405
Ala Ser Gly Asp Arg Thr Ile Lys Val Trp Asn Thr Ser Thr Cys
410 415 420
Glu Phe Val Arg Thr Leu Asn Gly His Lys Arg Gly Ile Ala Cys
425 430 435
Leu Gln Tyr Arg Asp Arg Leu Val Val Ser Gly Ser Ser Asp Asn
440 445 450
Thr Ile Arg Leu Trp Asp Ile Glu Cys Gly Ala Cys Leu Arg Val
455 460 465
Leu Glu Gly His Glu Glu Leu Val Arg Cys Ile Arg Phe Asp Asn
470 475 480
Lys Arg Ile Val Ser Gly Ala Tyr Asp Gly Lys Ile Lys Val Trp
485 490 495
Asp Leu Val Ala Ala Leu Asp Pro Arg Ala Pro Ala Gly Thr Leu
500 505 510
Cys Leu Arg Thr Leu Val Glu His Ser Gly Arg Val Phe Arg Leu
515 520 525
Gln Phe Asp Glu Phe Gln Ile Val Ser Ser Ser His Asp Asp Thr
530 535 540
Ile Leu Ile Trp Asp Phe Leu Asn Asp Pro Ala Ala Gln Ala Glu
545 550 555
Pro Pro Arg Ser Pro Ser Arg Thr Tyr Thr Tyr Ile Ser Arg
560 565

<210> 13
<211> 123
<212> PRT
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 3315936

<400> 13
Met Glu Phe Leu Glu Ile Gly Gly Ser Lys Pro Phe Arg Ser Tyr
1 5 10 15
Trp Glu Met Tyr Leu Ser Lys Gly Leu Leu Leu Ile Phe Val Val
20 25 30
Asp Ser Ala Asp His Ser Arg Leu Pro Glu Ala Lys Lys Tyr Leu
35 40 45
His Gln Leu Ile Ala Ala Asn Pro Val Leu Pro Leu Val Val Phe
50 55 60
Ala Asn Lys Gln Asp Leu Glu Ala Ala Tyr His Ile Thr Asp Ile
65 70 75
His Glu Ala Leu Ala Leu Ser Glu Val Gly Asn Asp Arg Lys Met
80 85 90
Phe Leu Phe Gly Thr Tyr Leu Thr Lys Asn Gly Ser Glu Ile Pro
95 100 105
Ser Thr Met Gln Asp Ala Lys Asp Leu Ile Ala Gln Leu Ala Ala
110 115 120
Asp Val Gln

<210> 14
<211> 1957
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 016108

<400> 14
attttgtca ctttctgtgt gaactaaagt gattcaatgt ctcttttgga ttgcttctgt 60
acttcaagaa cacaagttga atcactcaga cctgaaaaac agtctgaaac cagtatccat 120
caatacttgg ttgatgagcc aaccctttcc tggtcacgtc catccactag agccagtgaa 180
gtactatgtt ccaccaacgt ttctcactat gagctccaag tagaaatagg aagaggattt 240
gacaacttga cttctgtcca tcttgcacgg catactccca caggaacact ggtaactata 300
aaaattacaa atctggaaaa ctgcaatgaa gaacgcctga aagctttaca gaaagccgtg 360
attctatccc actttttccg gcatcccaat attacaactt attggacagt tttcactgtt 420
ggcagctggc tttgggttat ttctccattt atggcctatg gttcagcaag tcaactcttg 480
aggacctatt tccctgaagg aatgagtgaa actttaataa gaaacattct ctttggagcc 540
gtgagagggt tgaactatct gcaccaaaat ggctgtattc acaggagtat taaagccagc 600
catatcctca tttctggtga tggcctagtg accctctctg gcctgtccca tctgcatagt 660
ttggttaagc atggacagag gcatagggct gtgtatgatt tcccacagtt cagcacatca 720
gtgcagccgt ggttgagtcc agaactactg agacaggatt tacatgggtt atatgtgaag 780
tcagatattt acagtgttgg gatcacagca tgtgaattag ccagtgggca ggtgcctttc 840
caggacatgc atagaactca gatgctgtta cagaaactga aaggtcctcc ttatagccca 900
ttggatatca gtattttccc tcaatcagaa tccagaatga aaaattccca gtcaggtgta 960
gactctggga ttggagaaag tgtgcttgtc tccagtggaa ctcacacagt aaatagtgac 1020
cgattacaca caccatcctc aaaaactttc tctcctgcct tctttagctt ggtacagctc 1080
tgtttgcaac aagatcctga gaaaaggcca tcagcaagca gtttattgtc ccatgttttc 1140
ttcaaacaga tgaaagaaga aagccaggat tcaatacttt cactgttgcc tcctgcttat 1200
aacaagccat caatatcatt gcctccagtg ttaccttgga ctgagccaga atgtgatttt 1260
cctgatgaaa aagactcata ctgggaattc tagggctgcc aaatcatttt atgtcctata 1320

```
tacttgacac tttctccttg ctgcttttc ttctgtattt ctaggtacaa ataccagaat 1380
tatacttgaa aatacagttg gtgcactgga gaatctatta tttaaaacca ctctgttcaa 1440
aggggcacca gtttgtagtc cctctgtttc gcacagagta ctatgacaag gaaacatcag 1500
aattactaat ctagctagtg tcatttattc tggaattttt ttctaagctg tgactaactc 1560
tttttatctc tcaatataat ttttgagcca gttaattttt ttcagtattt tgctgtccct 1620
tgggaatggg ccctcagagg acagtgcttc caagtacatc ttctcccaga ttctctggcc 1680
tttttaatga gctattgtta aaccaacagg ctagtttatc ttacatcaga cccttttctg 1740
gtagagggaa aatgtttgtg ctttcccttt ttcttctgtt aatacttatg gtaacaccta 1800
actgagcctc actcacatta aatgattcac ttgaaatata tacagaaatt gtaatttgct 1860
ttttttaaa aaaggggggct aaagtaacac tttcctactt atgtaaatta tagatcctaa 1920
attcacgcac cccgtgggag ctcaataaag atttact 1957
```

<210> 15
<211> 2545
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 640521

<400> 15

```
gagccgagct gggggcgcag agcgcgggag gcggcggcgg cgcggaccca gtcacccagg 60
ctgcagtgca gtggtgcgat ctcggctcag tcattgcaac cttcacctcc cggattcaag 120
tgattctcct gcctcagcct cccgagtagc tgggattaca ggtgcccacc accatgccca 180
ggtggctccg ctgccggatg ggagtgcccc agtgtgctgg atgaagctgg cgcatgcacc 240
atgtcatcat gtgtctctag ccagcccagc agcaaccggg ccgcccccca ggatgagctg 300
gggggcaggg gcagcagcag cagcgaaagc cagaagccct gtgaggccct gcggggcctc 360
tcatccttga gcatccacct gggcatggag tccttcattg tggtcaccga gtgtgagccg 420
ggctgtgctg tggacctcgg cttggcgcgg gaccggcccc tggaggccga tggccaagag 480
gtcccccttg actcctccgg gtcccaggcc cggccccacc tctccggtcg caagctgtct 540
ctgcaagagc ggtcccaggg tgggctggca gccggtggca gcctggacat gaacggacgc 600
tgcatctgcc cgtccctgcc ctactcaccc gtcagctccc cgcagtcctc gcctcggctg 660
ccccggcggc cgacagtgga gtctcaccac gtctccatca cgggtatgca ggactgtgtg 720
cagctgaatc agtataccct gaaggatgaa attggaaagg gctcctatgg tgtcgtcaag 780
ttggcctaca atgaaaatga caatacctac tatgcaatga aggtgctgtc caaaaagaag 840
ctgatccggc aggccggctt tccacgtcgc cctccacccc gaggcacccg gccagctcct 900
ggaggctgca tccagcccag ggccccatt gagcaggtgt accaggaaat tgccatcctc 960
aagaagctgg accaccccaa tgtggtgaag ctggtggagg tcctggatga ccccaatgag 1020
gaccatctgt acatggtgtt cgaactggtc aaccaagggc ccgtgatgga agtgcccacc 1080
ctcaaaccac tctctgaaga ccaggcccgt ttctacttcc aggatctgat caaaggcatc 1140
gagtacttac actaccagaa gatcatccac cgtgacatca aaccttccaa cctcctggtc 1200
ggagaagatg ggcacatcaa gatcgctgac tttggtgtga gcaatgaatt caagggcagt 1260
gacgcgctcc tctccaacac cgtgggcacg cccgccttca tggcacccga gtcgctctct 1320
gagacccgca agatcttctc tgggaaggcc ttggatgttt gggctatggg tgtgacacta 1380
tactgctttg tctttggcca gtgcccattc atggacgagc ggatcatgtg tttacacagt 1440
aagatcaaga gtcaggccct ggaatttcca gaccagcccg acatagctga ggacttgaag 1500
gacctgatca cccgtatgct ggacaagaac cccgagtcga ggatcgtggt gccggaaatc 1560
aagctgcacc cctgggtcac gaggcatggg gcggagccgt tgccgtcgga ggatgagaac 1620
tgcacgctgg tcgaagtgac tgaagaggag gtcgagaact cagtcaaaca cattcccagc 1680
ttggcaaccg tgatcctggt gaagaccatg atacgtaaac gctcctttgg gaacccattc 1740
gagggcagcc ggcgggagga acgctcactg tcagcgcctg gaaacttgct cacgaagcaa 1800
ggcagcgaag acaacctcca gggcaccgac ccgccccccg tgggggagga ggaagtgctc 1860
ttgtgagagg cagtccctgc gtggaaagtt gctgggcccc cgcccccggc tcccccgcac 1920
gcatgcatcc actgcggccg gaggaggcca tggagcccga gtagctgcct ggatcgctcg 1980
acctcgcatg cgcgccgcgt cgcctctggg ggctgctgc accgcgtttc catagcagca 2040
tgtcctacgg aaacccagca cgtgtgtaga gcctcgatcg tcatctctgg ttatttgttt 2100
tttcctttgt tgttttaaag ggacaaaaa aaaaaaaaaa aaggacttga ctccatgacg 2160
tcgaccgtgg ccgctggctg gctggacagg cgggtgtgag gagttgcaga cccaaaccca 2220
cgtgcatttt gggacaattg ctttttaaaa cgtttttatg ccaaaaatcc ttcattgtga 2280
ttttcagaac cacgtcagat ataccaagtg actgtgtgtg gggtttgaca actgtggaaa 2340
ggcgagcaga aaactccggc ggtctgaggc catggaggtg gttgctgcat ttgagaggga 2400
gtagggggct agatgtggct cctagtgcaa accggaaacc atggcacctt ccagagccgt 2460
```

ggtctcaagg agtcagagca gggctggccc tcagtagctg cagggagctt tgatggcaac 2520
ttattttgtt aagaagggtt ttttt 2545

<210> 16
<211> 3034
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1250171

<400> 16
tcctgagtct cgaggaggcc gcgggagccc gccggcggtg gcgcggcgga gacccggctg 60
gtataacaag aggattgcct gatccagcca agatgcagag cacttctaat catctgtggc 120
ttttatctga tattttaggc caaggagcta ctgcaaacgt ctttcgtgga agacataaga 180
aaactggtga tttatttgct atcaaagtat ttaataacat aagcttcctt cgtccagtgg 240
atgttcaaat gagagaattt gaagtgttga aaaaactcaa tcacaaaaat attgtcaaat 300
tatttgctat tgaagaggag acaacaacaa gacataaagt acttattatg gaattttgtc 360
catgtgggag tttatacact gttttagaag aaccttctaa tgcctatgga ctaccagaat 420
ctgaattctt aattgttttg cgagatgtgg tgggtggaat gaatcatcta cgagagaatg 480
gtatagtgca ccgtgatatc aagccaggaa atatcatgcg tgttataggg gaagatggac 540
agtctgtgta caaactcaca gattttggtg cagctagaga attagaagat gatgagcagt 600
ttgtttctct gtatggcaca gaagaatatt tgcaccctga tatgtatgag agagcagtgc 660
taagaaaaga tcatcagaag aaatatggag caacagttga tctttggagc attggggtaa 720
cattttacca tgcagctact ggatcactgc catttagacc ctttgaaggg cctcgtagga 780
ataaagaagt gatgtataaa ataattacag gaaagccttc tggtgcaata tctggagtac 840
agaaagcaga aaatggacca attgactgga gtggagacat gcctgtttct tgcagtcttt 900
ctcggggtct tcaggttcta cttacccctg ttcttgcaaa catccttgaa gcagatcagg 960
aaaagtgttg ggtttttgac cagtttttg cagaaactag tgatatactt caccgaatgg 1020
taattcatgt tttttcgcta caacaaatga cagctcataa gatttatata catagctata 1080
atactgctac tatatttcat gaactggtat ataaacaaac caaaattatt tcttcaaatc 1140
aagaacttat ctacgaaggg cgacgcttag tcttagaacc tggaaggctg gcacaacatt 1200
tccctaaaac tactgaggaa aaccctatat ttgtagtaag ccgggaacct ctgaatacca 1260
taggattaat atatgaaaaa atttccctcc ctaaagtaca tccacgttat gatttagacg 1320
gggatgctag catggctaag gcaataacag ggggttgtgtg ttatgcctgc agaattgcca 1380
gtaccttact gctttatcag gaattaatgc gaaaggggat acgatggctg attgaattaa 1440
ttaaagatga ttacaatgaa actgttcaca aaaagacaga agttgtgatc acattggatt 1500
tctgtatcag aaacattgaa aaaactgtga aagtatatga aaagttgatg aagatcaacc 1560
tggaagcggc agagttaggt gaaatttcag acatacacac caaattgttg agactttcca 1620
gttctcaggg aacaatagaa accagtcttc aggatatcga cagcagatta tctccaggtg 1680
gatcactggc agacgcatgg gcacatcaag aaggcactca tccgaaagac agaaatgtag 1740
aaaaactaca agtcctgtta aattgcatga cagagattta ctatcagttc aaaaaagaca 1800
aagcagaacg tagattagct tataatgaag aacaaatcca caaatttgat aagcaaaaac 1860
tgtattacca tgccacaaaa gctatgacgc actttacaga tgaatgtgtt aaaaagtatg 1920
aggcattttt gaataagtca gaagaatgga taagaaagat gcttcatctt aggaaacagt 1980
tattatcgct gactaatcag tgttttgata ttgaagaaga agtatcaaaa tatcaagaat 2040
atactaatga gttacaagaa actctgcctc agaaaatgtt tacagcttcc agtggaatca 2100
aacataccat gaccccaatt tatccaagtt ctaacacatt agtagaaatg actcttggta 2160
tgaagaaatt aaaggaagag atggaagggg tggttaaaga acttgctgaa aataaccaca 2220
ttttagaaag gtttggctct ttaaccatgg atggtggcct tcgcaacgtt gactgtcttt 2280
agctttctaa tagaagttta agaaaagttt ccgtttgcac aagaaaataa cgcttgggca 2340
ttaaatgaat gcctttatag atagtcactt gtttctacaa ttcagtattt gatgtggtcg 2400
tgtaaatatg tacaatattg taaatacata aaaaatatac aaatttttgg ctgctgtgaa 2460
gatgtaattt tatcttttaa catttataat tatatgagga aatttgacct cagtgatcac 2520
gagaagaaag ccatgaccga ccaatatgtt gacatactga tcctctactc tgagtggggc 2580
taaataagtt attttctctg accgcctact ggaaatattt ttaagtggaa ccaaaatagg 2640
catccttaca aatcaggaag actgacttga cacgtttgta aatggtagaa cggtggctac 2700
tgtgagtggg gagcagaacc gcaccactgt tatactggga taacaatttt tttgagaagg 2760
ataaagtggc attattttat tttacaaggt gcccagatcc cagttatcct tgtatccatg 2820
taatttcaga tgaattatta agcaaacatt ttaaagtgaa ttcattatta aaaactattc 2880
attttttcc tttggccata aatgtgtaat tgtcattaaa attctaaggt catttcaact 2940
gttttaagct gtatatttct ttaattctgc ttactatttc atggaaaaaa ataaatttct 3000

caattttaaa aaattttttt ataaaaaaaa aaaa 3034

<210> 17
<211> 1337
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1911587

<400> 17
gaaagctgtg ggaccatcct ggcaaccccg gtgtttggct gggttctagc gtaccgtctg 60
tgtggccggt gggggacctg cggtcggagt gggagggcca gtctgcaccc aagaggtgga 120
agaggacggg ctttaggctg gaagcgcctt agaggagcca tttttccagg tggggcccca 180
ggcagaggct ccgacaggga gcctggccat agtcgcgcac caggggaggt ggagcgcgtc 240
ccagacccga gcccccgacc tcagccaaac ccattccttc tgtccttgga ggccagaggg 300
gactctgagc atcggaaagc aggatgcctg gtttgctttt atgtgaacca acagagcttt 360
acaacatcct gaatcaggcc acaaaactct ccagattaac agaccccaac tatctctgtt 420
tattggatgt ccgttccaaa tgggagtatg acgaaagcca tgtgatcact gcccttcgag 480
tgaagaagaa aaataatgaa tatcttctcc cggagtctgt ggacctggag tgtgtgaagt 540
actgcgtggt gtatgataac aacagcagca ccctggagat actcttaaaa gatgatgatg 600
atgattcaga ctctgatggt gatggcaaag atcttgtgcc tcaagcagcc attgagtatg 660
gcaggatcct gacccgcctc acccaccacc ccgtctacat cctgaaaggg ggctatgagc 720
gcttctcagg cacgtaccac tttctccgga cccagaagat catctggatg cctcaggaac 780
tggatgcatt tcagccatac cccattgaaa tcgtgccagg gaaggtcttc gttggcaatt 840
tcagtcaagc ctgtgacccc aagattcaga aggacttgaa aatcaaagcc catgtcaatg 900
tctccatgga tacagggccc ttttttgcag gcgatgctga caggcttctg cacatccgga 960
tagaagattc cccggaagcc cagattcttc ccttcttacg ccacatgtgt cacttcattg 1020
aaattcacca tcaccttggc tctgtcattc tgatcttttc cacccagggt atcagccgca 1080
gttgtgccgc catcatagcc tacctcatgc atagtaacga gcagaccttg cagaggtcct 1140
gggcctatgt caagaagtgc aaaaacaaca tgtgtccaaa tcggggattg gtgagccagc 1200
tgctggaatg ggagaagact atccttggag attccatcac aaacatcatg gatccgctct 1260
actgatcttc tccgaggccc accgaagggt actgaagagc ctcacctggg ggcattttgt 1320
gggtggaggg ccagagt 1337

<210> 18
<211> 1639
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 2079081

<400> 18
gacaaaagcc agacacattt caacatgagg gacccactga cagattgtcc gtataataaa 60
gtatacaaga acctaaagga gttttctcaa aatggagaga atttctgcaa acaggtcaca 120
tctgttcttc agcaaagggc aaacctggaa attagctatg ccaaaggact tcagaaactg 180
gcaagcaagc tgagcaaagc attacagaac acgagaaaaa gttgtgttag cagtgcctgg 240
gcctgggcct cagagggaat gaaatccaca gcggacctgc atcaaaaact tggcaaagca 300
attgaattgg aagcaataaa accgacttat caagtcctaa atgtacaaga gaagaagaga 360
aaatcacttg acaatgaagt tgaaaagaca gcaaatcttg tcattagcaa ctggaatcag 420
caaattaagg ccaagaagaa attaatggtt agtaccaaga aacatgaagc acttttccag 480
cttgtagaaa gctccaagca atctatgact gagaaggaga agcggaagct cctcaataaa 540
ctgacaaaat caactgaaaa gttggaaaag gaagatgaaa attactacca aaaaaacatg 600
gcgggttatt ctaccagact gaaatgggaa aacacactag agaactgcta ccagagcatt 660
ctggagctgg agaaggaaag aattcaactt ttatgcaata acttaaacca gtacagccaa 720
catatttctc tttttggcca aaccctgacc acatgccaca cgcagattca ctgtgccatc 780
agcaagattg acattgaaaa agatatccag gctgtaatgg aagaaactgc aattttatct 840
acagaaaaca aatctgagtt cctgttaacg gattactttg aagaagatcc taacagtgca 900
atggataaag agagacgaaa gtctttacta aaaccaaaat tattgagact gcagagagac 960

attgaaaaag cctcaaaaga caaggaaggc ctggaacgaa tgcttaaaac gtactccagc 1020
acctcctcct tctctgatgc aaagagccag aaagacacag cagcgttaat ggatgagaac 1080
aatttgaaac tagacctttt ggaagcgaac tcctacaaac tgtcatcaat gttagcagaa 1140
cttgagcaaa gacctcaacc cagccatcct tgtagtaatt ccatcttcag gtggagggaa 1200
aaggagcata ctcatagcta tgtgaaaata tctcggcctt ttttaatgaa gagattagag 1260
aatattgtga gcaaggcatc ttctggtggg cagagcaatc caggttcttc aactccagcc 1320
cctggtgcag cccagctcag cagcagactt tgcaaggcct tgtattcttt tcaagccagg 1380
caagatgatg agttgaattt ggaaaagggt gacattgtga ttatacacga gaaaaaagaa 1440
gaaggatggt ggtttggatc tttgaatggg aaaaaaggcc attttcctgc cgcttatgtg 1500
gaggagttac cttcaaatgc tggcaacaca gctacaaagg cataaaacaa gactctgaac 1560
atactacctt cacactcggt aatcaacaat acagtgtggt tcaaataaga ataaagtgct 1620
cttaccttta aaaaaaaaa 1639

<210> 19
<211> 1504
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 2472655

<400> 19
cgaaatcgta ggacttccga aagcagcggt ggcgtttgct tcactgcttg gaagtgtgag 60
tgcgcgaaga tgcgaaaggt ggttttgatc accggggcta gcagtggcat tggcctggcc 120
ctctgcaagc ggctgctggc ggaagatgat gagcttcatc tgtgtttggc gtgcaggaac 180
atgagcaagg cagaagctgt ctgtgctgct ctgctggcct ctcaccccac tgctgaggtc 240
accattgtcc aggtggatgt cagcaacctg cagtcggtct tccgggcctc caaggaactt 300
aagcaaaggt ttcagagatt agactgtata tatctaaatg ctgggatcat gcctaatcca 360
caactaaata tcaaagcact ttttctttggc ctcttttcaa gaaaagtgat tcatatgttc 420
tccacagctg aaggcctgct gacccagggt gataagatca ctgctgatgg acttcaggag 480
gtgtttgaga ccaatgtctt tggccatttt atcctgattc gggaactgga gcctctcctc 540
tgtcacagtg acaatccatc tcagctcatc tggacatcat ctcgcagtgc aaggaaatct 600
aatttcagcc tcgaggactt ccagcacagc aaaggcaagg aaccctacag ctcttccaaa 660
tatgccactg accttttgag tgtggctttg aacaggaact tcaaccagca gggtctctat 720
tccaatgtgg cctgtccagg tacagcattg accaatttga catatggaat tctgcctccg 780
tttatatgga cgctgttgat gccggcaata ttgctacttc gctttttttgc aaatgcattc 840
actttgacac catataatgg aacagaagct ctggtatggc ttttccacca aaagcctgaa 900
tctctcaatc ctctgatcaa atatctgagt gccaccactg gctttggaag aaattacatt 960
atgacccaga agatggacct agatgaagac actgctgaaa aattttatca aaagttactg 1020
gaactggaaa agcacattag ggtcactatt caaaaaacag ataatcaggc caggctcagt 1080
ggctcatgcc tataattcca gcactttggg aggccaaggc agaaggatca cttgagacca 1140
ggagttcaag accagcctga gaaacatagt gagcccttgt ctctacaaaa agaaataaaa 1200
ataatagctg ggtgtggtgg catgcgcatg tagtcccagc tactcagaag gatgaggtgg 1260
gaggatctct tgaggctggg aggcagaggt tgcagtgagc tgagattgtg ccactgcact 1320
ccagcctggg tgacagcgag accctgtctc aaaatatgta tatatttaat atatatataa 1380
aaccagagct gacaatgaca ctctggaaca ttgcatacct tctgtacatt ctggggtaca 1440
tggatttcta ctgagttgga taatatgcat ttgtaataaa ctatgaacta tgaaaaaaaa 1500
aaaa 1504

<210> 20
<211> 3096
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 2948818

<400> 20
gggtgttctt ataacttaag ttcagttttt tctttcctgt gaggaagagg cagtttttta 60
aatgaagcca tctctgggga aatcgtattg attgttgtag ctaaatacgg aatttttaaa 120

gtctttagta tgttgaactg gaaatatagg acatgcgttt gagatctact gtgagttgca 180
tcataaatac aaaggactga agttataaaa gagaaaagag aagtttgctg ctaaaatgaa 240
tctgagcaat atggaatatt ttgtgccaca cacaaaaagg tactgaggat ttacccccca 300
aaaaaaattg tcaatgagaa ataaagctaa ctgatatcaa aaagcagagc ctgctctact 360
ggccatcatg cgtaaagggg tgctgaagga cccagagatt gccgatctat cctacaaaga 420
tgatcctgag gaactttta ttggtttgca tgaaattggg catggaagtt ttggagcagt 480
ttattttgct acaaatgctc acaccagtga ggtggtggca attaagaaga tgtcctatag 540
tgggaagcag acccatgaga aatggcaaga tattcttaag gaagttaaat ttttacgaca 600
attgaagcat cctaatacta ttgagtacaa aggctgttac ttgaaagaac acactgcttg 660
gttggtgatg gaatattgct taggctcagc ctctgattta ttagaagttc ataaaaaacc 720
acttcaggaa gtggagatcg ctgccattac tcatggagcc ttgcatggac tagcctacct 780
acattctcat gcattgattc atagggatat taaagcagga aatattcttc taacagagcc 840
aggtcaggta aaactagctg attttggatc tgcttcaatg gcttctcctg ccaactcctt 900
cgtgggcaca ccttactgga tggctccaga ggtgatctta gctatggatg aaggacagta 960
tgatgggaaa gttgatattt ggtcacttgg catcacttgt attgaattgg cggaacggaa 1020
gccgcccctt ttcaacatga atgcaatgag tgccttatat cacattgccc agaatgactc 1080
cccaacgtta cagtctaatg aatggacaga ctcctttagg agatttgttg attactgctt 1140
gcagaaaata cctcaggaaa ggccaacatc agcagaacta ttaaggcatg actttgttcg 1200
acgagaccgg ccactacgtg tcctcattga cctcatacag aggacaaaag atgcagttcg 1260
tgagctagat aacctacagt accgaaaaat gaaaaaaata cttttccaag agacacggaa 1320
tggacccttg aatgagtcac aggaggatga ggaagacagt gaacatggaa ccagcctgaa 1380
cagggaaatg gacagcctgg gcagcaacca ttccattcca agcatgtccg tgagcacagg 1440
cagccagagc agcagtgtga acagcatgca ggaagtcatg gacgagagca gttccgaact 1500
tgtcatgatg cacgatgacg aaagcacaat caattccagc tcctccgtcg tgcataagaa 1560
agatcatgta ttcataaggg atgaggcggg ccacggcgat cccaggcctg agccgcggcc 1620
tacccagtca gttcagagcc aggccctcca ctaccggaac agagagcgct ttgccacgat 1680
caaatcagca tctttggtta cacgacagat ccatgagcat gagcaggaga acgagttgcg 1740
ggaacagatg tcaggttata agcggatgcg gcgccagcac cagaagcagc tgatcgccct 1800
ggagaacaag ctgaaggctg agatggacga gcaccgcctc aagctacaga aggaggtgga 1860
gacgcatgcc aacaactcgt ccatcgagct ggagaagctg gccaagaagc aagtggctat 1920
catagaaaag gaggcaaagg tagctgcagc agatgagaag aagttccagc aacagatctt 1980
ggcccagcag aagaaagatt tgacaacttt cttagaaagt cagaagaagc agtataagat 2040
ttgtaaggaa aaaataaaag aggaaatgaa tgaggaccat agcacaccca agaaaagagaa 2100
gcaagagcgg atctccaaac ataaagagaa cttgcagcac acacaggctg aagaggaagc 2160
ccaccttctc actcaacaga gactgtacta cgacaaaaat tgtcgtttct tcaagcggaa 2220
aataatgatc aagcggcacg aggtggagca gcagaacatt cgggaggaac taaataaaaa 2280
gaggacccag aaggagatgg agcatgccat gctaatccgg cacgacgagt ccacccgaga 2340
gctagagtac aggcagctgc acacgttaca gaagctacgc atggatctga tccgtttaca 2400
gcaccagacg gaactggaaa accagctgga gtacaataag aggcgagaaa gagaactgca 2460
cagaaagcat gtcatggaac ttcggcaaca gccaaaaaac ttaaaggcca tggaaatgca 2520
aattaaaaaa cagtttcagg acacttgcaa agtacagacc aaacagtata aagcactcaa 2580
gaatcaccag ttggaagtta ctccaaagaa tgagcacaaa acaatcttaa agacactgaa 2640
agatgagcag acaagaaaac ttgccatttt ggcagagcag tatgaacaga gtataaatga 2700
aatgatggcc tctcaagcgt tacggctaga tgaggctcaa gaagcagaat gccaggcctt 2760
gaggctacag ctccagcagg aaatggagct gctcaacgcc taccagagca aaatcaagat 2820
gcaaacagag gcacaacatg aacgtgagct ccagaagcta gagcagagag tgtctctgcg 2880
cagagcacac cttgagcaga agattgaaga ggagctggct gcccttcaga aggaacgcag 2940
cgagagaata aagaacctat tggaaaggca agagcgagag attgaaactt ttgacatgga 3000
gagcctcaga atgggatttg ggaatttggt tacattagat tttcctaagg aggactacag 3060
atgagattaa attttttgcc atttacaaaa aaaaaa 3096

<210> 21
<211> 1527
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 054191

<400> 21
ctccgcttga ggagaagcgc caagtgcgca tggggacgct atagcaattc gtttgctgtc 60
cttcctctcc ttcgaagatg acaaggccta ccatcgtttc ttcctgcctt tgggccgtca 120

ggcagttggt tgggacccgc tccaaccctc ggttcttcct gcaatacagt ggatacaatt 180
tgtcatggct actctgagtg ttataggttc aagttcactt attgcctatg ctgtattcca 240
taatatacag aaatctccag agataagacc actttttat ctgagcttct gtgacctgct 300
cctgggactt tgctggctca cggagacact tctctatgga gcttcagtag caaataagga 360
catcatctgc tataacctac aagcagttgg acagatattc tacatttcct catttctcta 420
caccgtcaat tacatctggt atttgtacac agagctgagg atgaaacaca cccagagtgg 480
acagagcaca tctccactgg tgatagatta tacttgtcga gttggtcaaa tggcctttgt 540
tttctcaagc ctgatacctc tgctattgat gacacctgta ttctgtctgg gaaatactag 600
tgaatgtttc caaaacttca gtcagagcca caagtgtatc ttgatgcact caccaccatc 660
agccatggct gaacttccac cttctgccaa cacatctgtc tgtagcacac tttattttta 720
tggtatcgcc attttcctgg gcagctttgt actcagcctc cttaccatta tggtcttact 780
tatccgagcc cagacattgt ataagaagtt tgtgaagtca actggctttc tggggagtga 840
acagtgggca gtgattcaca ttgtggacca acgggtgcgc ttctacccag tggccttctt 900
ttgctgctgg ggcccagctg tcattctaat gatcataaag ctgactaagc cacaggacac 960
caagcttcac atggcccttt atgttctcca ggctctaacg gcaacatctc agggtctact 1020
caactgtgga gtatatggct ggacgcagca caaattccac caactaaagc aggaggctcg 1080
gcgtgatgca gatacccaga caccattatt atgctcacag aagagattct atagcagggg 1140
cttaaattca ctggaatcca ccctgacttt tcctgccagt acttctacca tttttgaaa 1200
ctacaatact ggaacatcca ggaactggag ttattctacg ctaatggatt ggaaagaatg 1260
ttgggaaagg acatcttaaa tcttttctaa ctatgcccta aactgcagaa ctcaaaggaa 1320
atatagtgcc attgttagta gtcattctag atgaattggg agtatctctc cagttattcc 1380
cagattcact agtgatcctt aaagtctcta ttcagggaga ggaagacact ttccatctca 1440
gagatagact cgtgttacct tgatggatat tggatttgtc taagtctctt ctagaaaaaa 1500
taaattctag attattaaaa aaaaaaa 1527

<210> 22
<211> 2948
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1403604

<400> 22
aaagaaaggt cagccgcaag cgaacttagc actggctaca ccctcctcaa ttctggttgg 60
cgagatgcgc tcttcccgga agtgacgcac aagtgccggc ggaaggggaa gtccaggagc 120
atgggtggtt tttttccccc taccgaggtc cgtgaggtgt gtgctaacca aggggcggct 180
cacaaccgtg acagactgcc attcctgagt ctcttctggc catgggcccc cggagccgtg 240
agcgtcgggc aggcgcggta cagaacacca acgacagcag cgccctcagc aagcgttccc 300
tggccgcgcg cgggtacgtg caggacccct ttgccgcgtt gctggttccg ggcgcggcgc 360
gccgcgcacc gctcattcac cgaggctact acgtccgcgc acgcgccgtg aggcactgcg 420
tgcgcgcttt tttggagcag attggcgcgc cccaggccgc gcttcgcgcg cagatcttgt 480
ctctcggcgc tggcttcgac tcgctctatt ttcgcttaaa aaccgcgggc cgcctggccc 540
gggctgcagt ctgggaggtg gattttccgg acgtggcgcg gcgcaaagca gaaaggattg 600
gagagacgcc agagctgtgc gcgttaaccg ggcctttcga gaggggggag cccgcgtccg 660
cgctgtgctt tgagagcgca gactactgca tcctgggtct ggacttgcgg cagctccagc 720
gagtggagga ggccctgggc gccgcggggc tcgacgcagc ctcacccact ctgctcctgg 780
ccgaggcggt gctgacctac ctcgagccgg agagtgccgc ggccctcatc gcctgggcag 840
cccagcgttt tcctaatgcc cttttcgtgg tctatgagca gatgaggcct caagacgcct 900
ttggccagtt catgctgcaa cattttcggc agctaaactc ccccctgcat ggcctggagc 960
gttttcctga cgtggaggcg cagcggcgcc gcttccttca agctggctgg accgcctgcg 1020
gtgccgtgga cataaatgaa ttctatcact gctttcttcc cgcagaagaa cgccggcggg 1080
tggaaaatat tgaacccttt gacgaatttg aggagtggca tctgaagtgc gcccattatt 1140
tcattctggc agcttctagg ggagacaccc tctcccacac cctagtgttt ccatcctcag 1200
aggcatttcc tcgcgtaaat cctgcttcgc cttcaggggt attccctgcc agcgtagtca 1260
gtagcgaggg ccaggtccca aacctgaaga gatatggcca cgcctctgtc ttcttgagcc 1320
cagacgttat tctcagtgca ggaggatttg gagagcagga ggggcggcac tgccgagtga 1380
gccagtttca cttgctctca agagattgtg actctgaatg gaaaggcagc caaataggca 1440
gttgtgggac tggagttcag tgggatggac gcctttatca caccatgaca agactctcag 1500
agagtcgggt tctggttctg ggagggagac tgtccccagt aagtccagcc ttggggggttc 1560
tccagcttca tttttttaag agtgaggata ataacactga ggacctgaaa gtgacaataa 1620
caaaggctgg ccgaaaggat gattccactt tgtgttgttg gcggcattca acaacagaag 1680

tgtcctgtca gaatcaggaa tatttgtttg tgtatggggg tcgaagcgtg gtggaacctg 1740
tactaagtga ctggcatttc ctccatgtag ggacaatggc ttgggtcagg atcccagtgg 1800
agggagaagt acctgaagcc cggcattctc acagtgcctg cacttggcaa gggggagccc 1860
ttattgctgg aggtctcggg gcttctgagg agccattgaa ctctgtgctc tttctgagac 1920
caatctcttg tggattcctc tgggagtcag tagacatcca gcctcccatt accccaaggt 1980
actcccacac agctcatgtg ctcaatggaa agctgttact ggttggaggg atctggattc 2040
attcctcctc atttcctgga gtgactgtga tcaatttgac tacaggattg agctctgagt 2100
atcagattga cacaacatat gtgccatggc cattaatgtt acacaaccat actagtatcc 2160
ttcttcctga agagcaacag ctcctgctcc ttggaggtgg tgggaactgc ttttcctttg 2220
gtacctactt caaccccat acagtcacat tagacctttc ttccttaagt gctgggcagt 2280
aaggactgga ctaatattca ggacccacta aagtagacaa taaagtttc cacaaatagg 2340
atgaccctct agctatagat actgccactc ctcctttccc catccttttt ttcccttagc 2400
actattcagt gcaaaaagtg aaaaaggttg gtaaaatagg taaaatacct agaaacaatc 2460
actacagaaa acagctgaag acagtggcca tgcagtccga gaggagtagt ggtctgcctc 2520
taattttcta atctaagttc gtttattgag ttacagtggt ctttagtaaa gtaaaacaat 2580
ttcccaatcc caggccttgt gatttgagat ggtaccttag aaaaagttac acgcagttcc 2640
gtggttgaat atatttgaga tggtacctta gaaaaagttt cacgcagatc cttggttgaa 2700
tatagttgag ggagcgtagt attgacaatt cttcatgtag gaaacctgaa atgaacacag 2760
tcacagtttg attaaaacat tgtcctgttt gttgcaacag aaaactcgga tagttttaac 2820
aacaggaaac acttgtagga cttcctttac caacatactt tttaaatgtt ttgctattgg 2880
ttccatattt atttagattt ataagtgtca ataaagcaaa cttttgatgc ctcaaaaaaa 2940
aaaaaaa 2948

<210> 23
<211> 1808
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1652936

<400> 23
gagagtatta aatgtgaatt gcctctgcct cgctttatga tgtcaaagaa tgatggtgag 60
atacgatttg ggaacccagc tgagctgcat ggaactaagg ttcagattcc atatttaact 120
acagaaaaaa attccttcaa aagaatggac gatgaggata aacaggaaac acagtctccc 180
acgatgtcac ccctcgcctc ccctccttct tccccgcctc attaccagag ggtgcccttg 240
agccatggct acagcaaact gcggagcagc gcagaacaga tgcatccagc accttatgaa 300
gctaggcagc cccttgtcca gcccgaggga tcctcatcag gggcccagg aaccaagccc 360
cttcggcatc aggcttccct tatccggtcc ttttcagtgg agagggaact acaggataac 420
agcagttacc ccgacgaacc ctggaggata acagaagaac agcgcgagta ctatgtcaat 480
cagttccgat cccttcagcc agacccaagc tctttcattt caggttctgt ggccaagaac 540
ttcttcacca aatcaaagct ttccattcca gaactctcct atatatggga gcttagtgat 600
gctgactgtg atggagccct gaccctgcct gagttctgtg ctgcgtttca tctcattgtg 660
gctcggaaga acggctaccc attgcctgag ggcctccctc caactctgca gccagaatac 720
ctgcaggcag cttttcctaa gcccaaatgg gactgtcaat tatttgattc ttattctgag 780
tcactgccgg caaatcaaca acctcgtgac ttgaatcgga tggagacatc tgttaaagac 840
atggctgacc ttcctgtccc taaccaggat gtaactagtg atgacaaaca agctttgaaa 900
agtactatca atgaagcctt accaaaggac gtgtctgagg atccagcaac tcccaaggat 960
tccaacagtc tcaaagcaag accaagatcc agatcttact ctagcacctc catagaagag 1020
gccatgaaaa ggggcgagga ccctcccacc ccgccacctc ggccacagaa aacccattcc 1080
agagcctcct ccttggatct gaataaagtc ttccagccca gtgtgccagc taccaagtca 1140
ggattgttac ccccaccacc tgcgctccct ccaagacctt gtccatcaca gtctgaacaa 1200
gtgtcggagg ccgagttact cccacagctg agcagagccc catcccaggc tgcagaaagt 1260
agtccagcaa agaaggatgt actgtattct cagccaccat caaagcccat tcgtaggaaa 1320
ttcagaccag aaaaccaagc tacagaaaac caagagcctt ccactgctgc aagtgggcca 1380
gcttctgcgg caaccatgaa accgcatcca acagtccaaa agcagtcttc caaacagaag 1440
aaggccattc aaactgctat ccgcaaaaat aaagaggcaa acgcagtgct ggctcggctg 1500
aacagtgagc tccagcagca gctcaaggag gttcatcaag aacgaattgc attggaaaac 1560
caattggaac aacttcgtcc ggtcactgtg ttgtgacccc cccatggttc aagtgacagt 1620
gggtgacctt gtctgccaag atctttcttt tgaatgtttt gaacccaact acttgtcata 1680
gatgtttgac tgtgtcaaaa gctgtgagca gcaaaatata atccatatga ccttttctct 1740
tgtatagact taaaaaaaaa aaaatagatc tttaattaag cggtcgcaag cttattccct 1800

ttagtgag 1808

<210> 24
<211> 1148
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 1710702

<400> 24
tgcgtacgtg cgtcgtctct atggtggcgg cggatttgga gggaccctac gaaccaggag 60
tcaggcgagc cgatctgggg ctgcaggatg ttccgctggg agcgctccat tcccctgcga 120
ggctcggccg ccgccctgtg caacaacctc agtgtgctgc agctgccggc tcgcaacctc 180
acgtattttg gcgtggttca tggaccaagc gcccagcttc tcagcgctgc tcctgagggt 240
gtgcccttgg cccagcgcca gctccacgct aaggagggtg ctggagtgag tcccccactt 300
atcactcagg tccactggtg tgtcctcccc ttccgagtgc tgctggtact cacctcacat 360
cgaggaatac agatgtacga gtccaatggc tacaccatgg tctactggca tgcactggac 420
tctggagatg cctccccagt acaggctgtg tttgcccggg gaattgctgc cagtggccac 480
ttcatctgtg tgggaacgtg gtcaggccgg gtgctggtgt ttgacatccc agcaaagggt 540
cccaacattg tactgagcga ggagctggct gggcaccaga tgccaatcac agacattgcc 600
accgagcctg cccagggaca ggattgtgtg gctgacatgg tgacggcaga tgactcaggc 660
ttgctgtgtg tctggcggtc agggccagaa ttcacattat tgacccgcat tccaggattt 720
ggagttccgt gcccctctgt gcagctgtgg caggggatca tagcagcagg ctatgggaac 780
ggacaagtgc atctatatga ggccactaca ggaaatctac atgtccagat caatgcccat 840
gcccgggcca tctgcgccct ggacctggct tctgaggtgg gcaagctact ctctgcaggt 900
gaggacacct ttgtgcatat ctggaagctg agcagaaacc cagagagtgg ctacattgag 960
gtggaacact gtcatggtga gtgtgtcgcc gacacccagc tgtgtggtgc tcgattttgt 1020
gattcctcag gcaactcctt tgctgtgact ggctatgacc ttgcggagat ccggagattc 1080
agcagtgtgt gagaagagca gccttccttt gtccctgtgg tattcataaa gtacccgctc 1140
cacccaaa 1148

<210> 25
<211> 2419
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 3239149

<400> 25
cggacgcgtg gggcctggc accaaagggg cggccccggc ggagagcgga cccagtggcc 60
tcggcgatta tggacccggc cgaggcggtg ctgcaagaga aggcactcaa gtttatgaat 120
tcctcagaga gagaagactg taataatggc gaaccccta ggaagataat accagagaag 180
aattcactta gacagacata caacagctgt gccagactct gcttaaacca agaaacagta 240
tgtttagcaa gcactgctat gaagactgag aattgtgtgg ccaaaacaaa acttgccaat 300
ggcacttcca gtatgattgt gcccaagcaa cggaaactct cagcaagcta tgaaaaggaa 360
aaggaactgt gtgtcaaata ctttgagcag tggtcagagt cagatcaagt ggaatttgtg 420
gaacatctta tatcccaaat gtgtcattac caacatgggc acataaactc gtatcttaaa 480
cctatgttgc agagagattt cataactgct ctgccagctc ggggattgga tcatattgct 540
gagaacattc tgtcatacct ggatgccaaa tcactatgtg ctgctgaact tgtgtgcaag 600
gaatggtacc gagtgacctc tgatggcatg ctgtggaaga agcttatcga gagaatggtc 660
aggacagatt ctctgtggag aggcctggca gaacgaagag gatggggaca gtatttattc 720
aaaaacaaac ctcctgacgg gaatgctcct cccaactctt tttatagagc actttatcct 780
aaaattatac aagacattga gacaatagaa tctaattgga gatgtggaag acatagttta 840
cagagaattc actgccgaag tgaaacaagc aaaggagttt actgtttaca gtatgatgat 900
cagaaaatag taagcggcct tcgagacaac acaatcaaga tctgggataa aaacacattg 960
gaatgcaagc gaattctcac aggccataca ggttcagtcc tctgtctcca gtatgatgag 1020
agagtgatca taacaggatc atcggattcc acggtcagag tgtgggatgt aaatacaggt 1080
gaaatgctaa acacgttgat tcaccattgt gaagcagttc tgcacttgcg tttcaataat 1140

```
ggcatgatgg tgacctgctc caaagatcgt tccattgctg tatgggatat ggcctcccca 1200
actgacatta ccctccggag ggtgctggtc ggacaccgag ctgctgtcaa tgttgtagac 1260
tttgatgaca agtacattgt ttctgcatct ggggatagaa ctataaaggt atggaacaca 1320
agtacttgtg aatttgtaag gaccttaaat ggacacaaac gaggcattgc ctgtttgcag 1380
tacagggaca ggctggtagt gagtggctca tctgacaaca ctatcagatt atgggacata 1440
gaatgtggtg catgtttacg agtgttagaa ggccatgagg aattggtgcg ttgtattcga 1500
tttgataaca agaggatagt cagtggggcc tatgatggaa aaattaaagt gtgggatctt 1560
gtggctgctt tggacccccg tgctcctgca gggacactct gtctacggac ccttgtggag 1620
cattccggaa gagtttttcg actacagttt gatgaattcc agattgtcag tagttcacat 1680
gatgacacaa tcctcatctg ggacttccta aatgatccag ctgcccaagc tgaacccccc 1740
cgttcccctt ctcgaacata cacctacatc tccagataaa taaccataca ctgacctcat 1800
acttgcccag gacccattaa agttgcggta tttaacgtat ctgccaatac caggatgagc 1860
aacaacagta acaatcaaac tactgcccag tttccctgga ctagccgagg agcagggctt 1920
tgagactcct gttgggacac agttggtctg cagtcggccc aggacggtct actcagcaca 1980
actgactgct tcagtgctgc tatcagaaga tgtcttctat cttttgtgaa tgattggaac 2040
ttttaaacct cccctcctct cctcctttca cctctgcacc tagttttttc ccattggttc 2100
cagacaaagg tgacttataa atatatttag gtgttttngc ccaggaatct ctcttgcttt 2160
ggccattaag gcaggaggaa ctaggtttcc cctgtatagg gcctgcgggg ggagaggacc 2220
ccactctagg gggtaggggg gggggtgnca gctttcaagg cccaggggcc ccaggtgtct 2280
tccccggtta actgcagggg atgtccagga ccgggggggc tacgagcaag gcccggcccc 2340
ataggtctag gggagggggа cagagttccc ctcgtaatag ggctcggggg agggcaggga 2400
aagggaaaca caggatttg                                              2419
```

<210> 26
<211> 746
<212> DNA
<213> Homo sapiens

<220>
<221> misc-feature
<223> Incyte Clone 3315936

<400> 26

```
atttaatatg actcactata gggaatttgg ccctcgagct agagattcgg gcacgagggg   60
ttgcttagac tgcggcccac gtggaaggct cttagccacc ctgcctggcc cgaggagaaa 120
aacaagcaaa tcctagtgct gggcctggat ggagcaggaa aaaccagtgt cctgcactct 180
ctagcttcaa acagagtcca gcacagtgtg gcacccaccc aaggtttcca tgcagtttgc 240
atcaacactg aagacagcca gatggagttc ctggagattg gtggcagtaa accttttcgg 300
tcctactggg aaatgtacct atccaaggga ttgctgctga tctttgtggt ggattcagca 360
gatcacagcc gattacctga agccaagaaa taccttcatc agctaattgc agcaaaccca 420
gtacttcctc tggttgtgtt tgcaaacaaa caggatcttg aagcagccta tcacattaca 480
gatatccatg aagctttggc attatctgaa gtgggaaatg acaggaagat gttcttgttt 540
ggaacctacc tgactaagaa tggctcagag ataccctcca ccatgcaaga tgccaaagac 600
ttgattgcac agctggctgc agatgtgcag tgaccaggac tcagcccact gtgcggctca 660
cgactgagat gtcatcagtg ttgaatggca ggcttgaagc caaaggtttc cacctcaaat 720
aaaaattaag ccatttccta ttaaaa                                      746
```